# Adult-attained height, early life energy restriction, genetic variation, and cancer risk

# Adult-attained Height, Early Life Energy Restriction, Genetic Variation, And Cancer Risk

*Rachel Johanna Jacoba Elands*

# Adult-attained height, early life energy restriction, genetic variation, and cancer risk

**Rachel Elands**

**Adult-attained height, early life energy restriction, genetic variation, and cancer risk**

ISBN: 978-94-92801-42-5
Cover: By Anne-Marij van Rooije, Groningen and Rachel Elands
Layout and production: Guus Gijben



The work presented in this thesis was performed within GROW, School for Oncology and Developmental Biology at the Department of Epidemiology, Maastricht University, the Netherlands. The studies described within this thesis are part of the Netherlands Cohort Study on Diet and Cancer and were funded by a grant [RFA 2012/618] obtained from Wereld Kanker Onderzoek Fonds (WCRF NL), as part of the World Cancer Research Fund International grant program.


Printing and dissemination of this thesis was financially supported by the Department of Epidemiology at Maastricht University and Agena Bioscience.

# Adult-attained height, early life energy restriction, genetic variation, and cancer risk

DISSERTATION

To obtain the degree of Doctor at Maastricht University, on the authority of the Rector Magnificus, Prof. Dr. Rianne M. Letschert in accordance with the decision of the Board of Deans, to be defended in public on Tuesday, 3rd of July 2018 at 14:00 hours

by

**Rachel Johanna Jacoba Elands**
Born 6th of August 1987,'s-Hertogenbosch, the Netherlands

**Promotores**
Prof. Dr. Ir. M.P. Weijenberg
Prof. Dr. Ir. P.A. van den Brandt

**Co-promotor**
Dr. C.C.J.M. Simons

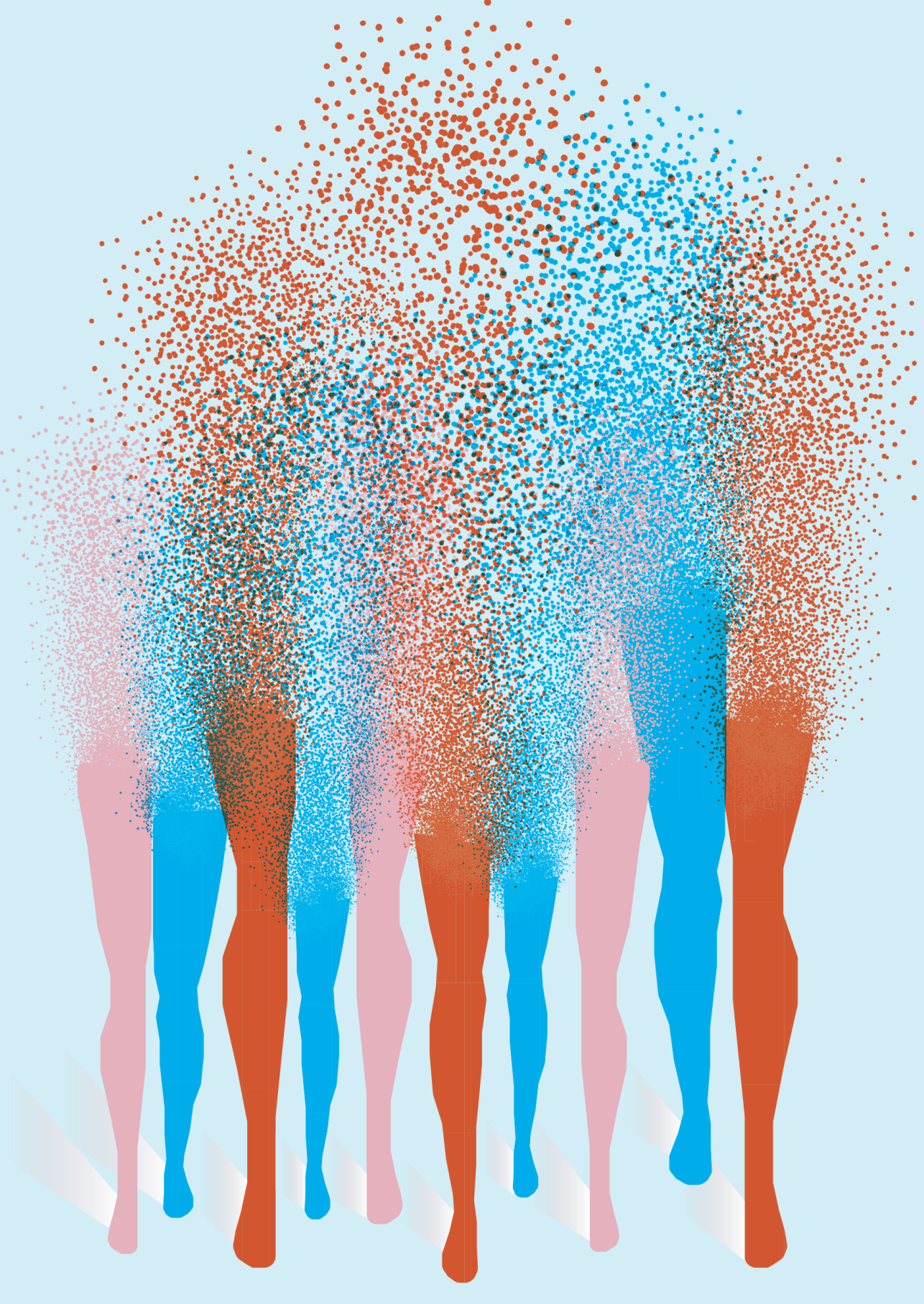
**Assessment Committee**
1. Prof. Dr. F.J. van Schooten *(Chair)*
2. Dr. F.J.B. van Duijnhoven *(Wageningen University)*
3. Prof. Dr. M. van Engeland
4. Prof. Dr. C.H. van Gils *(Julius Centrum, UMC Utrecht)*
5. Dr. A.C. Voogd

The work presented in this thesis was performed within GROW, School for Oncology and Developmental Biology at Maastricht University. This work was supported by a grant [RFA 2012/618] obtained from Wereld Kanker Onderzoek Fonds (WCRF NL), as part of the World Cancer Research Fund International grant program.

Financial support by Agena Bioscience for the publication of this thesis is acknowledged.

*Aan mijn opa en beide oma's*
*voor al hun wijsheid*

# Table of Contents



Chapter 1

# General introduction

## 1. Adult-attained height and cancer: the epidemiological evidence

Over the past decades, human height has increased in most industrialized countries.[1,2] This trend has been attributed to improvements in nutritional, hygienic, economic, and health status since this upward trend has occurred to a greater extent in developed than in developing countries.[3,4] The Dutch population has been well studied and is currently the tallest in the world[5] with women measuring, on average, 169.2 cm and men 183.2 cm in 2016[6] (source: CBS, Statistics Netherlands). In the Netherlands, people have been getting taller, as documented since 1858,[1,7] with a height gain of 21 cm over a period of 140 years being reported in several studies.[1,7,8] As in most Northern European countries,[1,2,5,9-11] this upward trend has been slowing down significantly in the Netherlands in recent years.[1,5] Still, average heights are continuing to increase in most countries,[4,12-15] particularly in those undergoing rapid socioeconomic improvement.[9,16] This secular trend of lower prevalence of infectious disease experienced during early life and improved nutrition may influence adult-attained height and may also contribute to the development of chronic disease in later life.[17,18] Because adult-attained height is known to be influenced by these early life factors, it has been widely used as a surrogate measure for early life exposures in epidemiologic studies.[17] Although numerous but sometimes conflicting findings in the literature have associated adult-attained height with morbidities and mortality,[17-20] the association between adult-attained height and increased cancer risk is relatively consistent.[21-26] Taller adults have higher rates of several types of cancer compared to shorter adults.[21-26] The World Cancer Research Fund has released several reports including the summary of the evidence for the association between adult-attained height and the risk of site-specific cancer. The World Cancer Research Fund concluded that there is strong evidence that height is a convincing risk factor for colorectal cancer in men and women,[21] and for breast cancer (both premenopausal and postmenopausal) [22] and ovarian cancer[27] (**Table 1**). There is also strong evidence that height is probably a risk factor for kidney cancer in men and women,[23] pancreatic cancer in men and women[24] and prostate cancer[25]. For other cancer sites, the evidence has been classified as limited *i.e.,* limited – suggestive (endometrial cancer and thyroid cancer) or limited - no conclusion (oesophagus cancer, gallbladder cancer, stomach cancer and, cervical cancer).[28,29] Understanding the height-cancer association may offer insights into the future cancer burden and allows anticipation of future health needs.

**Table 1. Associations between adult-attained height and site-specific cancer risk**

| Cancer site | Strength of the association RR per 5 cm (95% CI) | | Level of evidence for an association | | CUP** publishing date |
|---|---|---|---|---|---|
| | Male | Female | Male | Female | |
| Colorectal cancer | 1.04 (1.03-1.05) | 1.06 (1.02-1.09) | Convincing | Convincing | 2017 |
| Postmenopausal breast cancer | * | 1.09 (1.07-1.11) | * | Convincing | 2017 |
| Premenopausal breast cancer | * | 1.06 (1.02- 1.11) | * | Convincing | 2017 |
| Ovarian cancer | * | 1.08 (1.05-1.10) | * | Convincing | 2014 |
| Kidney cancer | 1.10 (1.06-1.13) | 1.10 (1.07-1.14) | Probable | Probable | 2015 |
| Pancreatic cancer | 1.07 (1.01-1.14) | 1.07 (0.99-1.15) | Probable | Probable | 2012 |
| Prostate cancer | 1.03-1.05) | * | Probable | * | 2014 |

*Information unavailable

**Continuous Update Project, for each cancer endpoint published on a regular basis by the World Cancer Research Fund.

## 2. Mechanisms linking adult-attained height to cancer

Although there is a broad understanding of how early life environmental[30] and genetic[31] processes contribute to adult-attained height, there is still little evidence on how these factors might link to cancer risk.[32] As it has been observed that height increases the risk of a number of different types of cancer and because risk estimates are very similar across different cancers and in different populations, a common mechanism might be at play.[33] This mechanism probably relates to cell growth, for which a person's adult-attained height may be regarded a consistent marker of mechanisms influencing increased cell growth. Cell growth and proliferation leads to a larger number of cells and taller individuals may be at an increased risk of cancer simply because they have more cells in their body and, therefore, have a greater chance of acquiring cancerous cell changes in one of these cells as compared to shorter people.[34-37] Cell growth can be influenced by (hormonal) growth factors, which, in turn, can be influenced by environmental exposures early in life (*e.g*. early life energy restriction) and genetic variation relevant to (hormonal) growth factors (**Figure 1**). By doing so, early life environmental exposures and genetic factors may influence both adult-attained height and cancer risk later in life.

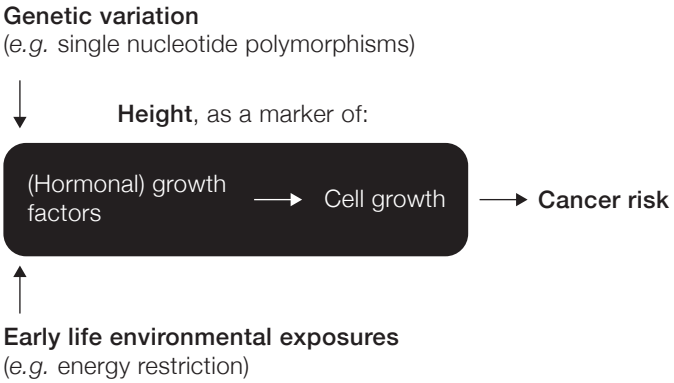



**Figure 1. Schematic overview of factors that may link adult-attained height to cancer risk.** Adult-attained height is possibly a marker for cell growth that can be influenced by (hormonal) growth factors, which, in turn, can be influenced by genetic variation and early life environmental exposures.

## 2.1 Early life environmental exposures: early life energy restriction

Adult-attained height is informative as an indicator of energy balance and is a reflection of environmental and nutritional exposures that start during intrauterine development and continue in the next 20 years of life.[38] Major determinants of height are modifiable early life exposures, *e.g.* caloric intake,[39,40] macronutrient composition,[41,42] childhood infections,[39-41,43] and socioeconomic status,[39,40,44] which may also influence cancer risk later in life. Even though some of these early life exposures may be transient in nature, the effect on cell growth and thereby on adult-attained height and carcinogenesis may be lasting. As most cancers have a long latency period, the influence of nutritional factors during childhood and adolescence, may play an important role. For instance, childhood obesity, which is on the positive side of the energy balance spectrum, is an established risk factor for many cancers.[45,46] In this thesis, we focus specifically on postnatal energy restriction. It is interesting to explore how energy restriction, which is on the other end of the energy balance spectrum, is related to cancer risk. Energy restriction early in life is insightful to investigate in relation to cancer risk in human observational studies given that it is an extreme exposure and also because it marks a fairly defined period early in life compared, for instance, to childhood obesity. Evidence from animal experimental models has consistently shown

that energy restriction reduces cancer risk at multiple sites.[47,48] Interestingly, timing of exposure to energy restriction may be important as energy restriction occurring early in life appears to be especially protective against cancer risk in animal models.[49,50] Other aspects of energy restriction, such as the duration and the severity of exposure, may determine whether energy restriction is associated with an increased or decreased risk in animal models.[51,52] Studying the long-term consequences of early life energy restriction on cancer risk in human experimental studies is, of course, not ethical. Therefore, evidence for an association of early life energy restriction with cancer risk in humans can only be derived from observational studies. In these observational studies, energy restriction exposure is mostly early in life and due to economic recession or war.[53-66] Generally, associations between early life energy restriction and cancer risk in human observational studies have been inconclusive. Given the evidence from animal models on the energy restriction-cancer association, the timing, duration, and severity of early life energy restriction may also be relevant when studying cancer risk in human observational studies.

## 2.2 Genetic factors: genetic variation

Depending on the nutritional contrast in a population, it has been reported that 69-98% of the variation in height is, among other inherited factors, determined by genetic variation.[67-71] The most common type of genetic variation is the single nucleotide polymorphism (SNP).[72] SNPs are often being studied in observational studies within a genome-wide set of other common SNPs, known as genome-wide association studies (GWAS), in which each SNP is present to some appreciable degree within a population (*i.e.* > 5%).[73] GWAS have investigated common SNPs in relation to several traits and diseases, such as adult-attained height[74] and cancer risk.[75-79] Nowadays, a large number of SNPs have been identified to be associated with either adult-attained height[80] or cancer risk.[81] Nevertheless, GWAS SNPs are not necessarily causal variants but may simply be (highly) correlated to the causal variants.[82,83] Therefore, it might be more informative to look at clusters of GWAS SNPs occurring in regions of low recombination, likely pinpointing spots in the genome that harbor causal variation. Recently, it has been demonstrated that SNPs associated with complex traits or diseases tend to co-segregate in certain genomic regions, encompassing one or more genes.[84] This cluster approach may also be fruitful when one is interested in the mechanisms underlying the association between two phenotypes, such as adult-attained height and cancer risk. Clusters may be identified harboring SNPs for both phenotypes. Genes in these clusters, in turn, may collectively point to

shared biological processes, such as cell growth, explaining the link between adult-attained height and cancer risk. A recent meta-analysis of GWAS on height identified SNPs in genes that act together in biologically relevant pathways[31] that not only appear to be important for normal tissue development but have also been implicated in cancer development.[85] Therefore, it may be interesting to study genomic regions, in which SNPs cluster together that are associated with adult-attained height and those that are associated with cancer risk. In this way, GWAS SNPs can be viewed as time-independent markers of biological processes (*e.g.* exposure to (hormonal) growth factors) that may link adult-attained height to cancer risk [71,86] (**Figure 2**).

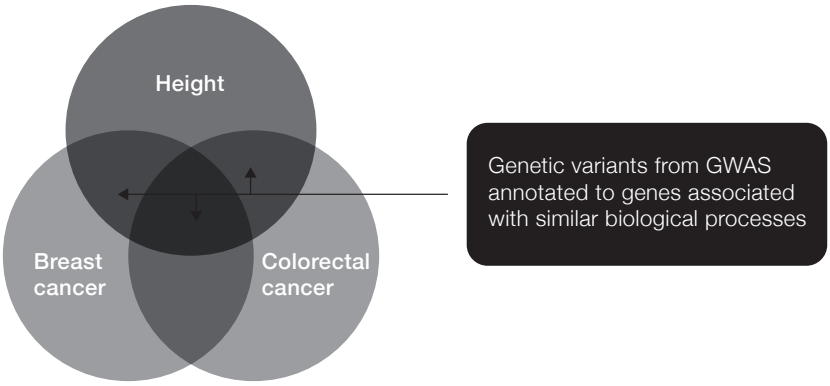



**Figure 2. Genetic variants from GWAS on height, breast cancer risk, and/or colorectal cancer risk can be used as markers of pathway involvement underpinning the height-cancer association.** The set of GWAS SNPs that tend to co-segregate in certain genomic regions, encompassing one or more genes, may collectively point to shared biological processes, such as cell growth, explaining the link between adult-attained height and cancer risk.

# 3. Thesis objectives, study design and outline

## 3.1 Objectives

Our main objective was to gain a deeper understanding of how height influences cancer risk later in life. As outlined in this chapter, the height-cancer risk association is consistent in the literature, but plausible mechanisms remain to be further

elucidated. Consistency and the presence of plausible underlying mechanisms are two of the Bradford-Hill criteria for causality.[87-89] We aim to specifically contribute to the evidence for plausible mechanisms. We do so by studying how early life energy restriction influences both adult-attained height and cancer risk. Early life determinants of growth are presumably associated with cancer risk later in life in a direction as expected based on analogy with the height-cancer association. Analogy of similar factors is another Bradford-Hill criterion for causality. We also specifically studied height in relation to the risk of postmenopausal breast cancer by hormone receptor subtypes. Breast cancer includes hormone-sensitive tumors, which may be susceptible to hormonal growth factors influencing adult-attained height and cancer risk. Finding an association between height and hormone receptor-positive breast cancer indicates involvement of hormonal growth factors as plausible mechanisms. Furthermore, a molecular epidemiological approach has been applied to identify genetic variants and genes that may link height to cancer risk. We focused on postmenopausal breast cancer and colorectal cancer. Both types of cancer share a subset of risk factors[90] and height has been identified as a convincing risk factor for these cancers according to the World Cancer Research Fund. Germline genetic variants are useful as markers of shared mechanisms between height and cancer because these fulfill the Bradford-Hill criterion of temporality.

## 3.2 Study design and outline

This thesis begins with a systematic review and meta-analysis in which the evidence on the association between early life energy restriction and cancer risk in humans is investigated (Chapter 2). Furthermore, specific aspects of early life energy restriction are discussed, such as the timing, duration, and severity of exposure, which may determine whether exposure is associated with an increased or decreased risk of cancer. Chapter 3 presents associations of adult-attained height and early life energy restriction with postmenopausal breast cancer risk according to estrogen and progesterone receptor status investigated within the Netherlands Cohort Study on Diet and Cancer. The Netherlands Cohort Study is a prospective cohort study among 120,852 men and women and has data available on adult-attained height and the rather unique exposure of early life energy restriction. In Chapter 4, a SNP selection approach is presented, which was developed to select SNPs to identify mechanisms linking height to postmenopausal breast- and colorectal cancer risk. Finally, Chapter 5 concludes the thesis with a discussion of the main findings as well as the challenges of and insights offered by (molecular) epidemiology. In addition, implications and suggestions for future research are discussed that allow us to gain a deeper understanding of how height influences cancer risk later in life.

# References


1 Fredriks, A. M. *et al.* Continuing positive secular growth change in The Netherlands 1955-1997. *Pediatr Res* **47**, 316-323 (2000).

2 Gohlke, B. & Woelfle, J. Growth and puberty in German children: is there still a positive secular trend? *Dtsch Arztebl Int* **106**, 377-382, doi:10.3238/arztebl.2009.0377 (2009).

3 Hauspie, R. C., Vercauteren, M. & Susanne, C. Secular changes in growth. *Horm Res* **45**, 8-17 (1996).

4 Hatton, T. J. How have Europeans grown so tall? *Oxford Econ Pap* **66**, 349-372, doi:10.1093/oep/gpt030 (2014).

5 Schonbeck, Y. *et al.* The world's tallest nation has stopped growing taller: the height of Dutch children from 1955 to 2009. *Pediatr Res* **73**, 371-377, doi:10.1038/pr.2012.189 (2013).

6 CBS. Specified height in the Netherlands in 2016 by age group (aged 20-30 years) and gender. Statistics Netherlands. https://opendata.cbs.nl/statline Accessed 28th November 2017.

7 Tassenaar, J. W. D. a. V. Paradoxes of Modernization and Material Well-Being in the Netherlands during the Nineteenth Century. p. 331 - 378 (1997).

8 de Beer, H. Observations on the history of Dutch physical stature from the late-Middle Ages to the present. *Economics & Human Biology* **2**, 45-55, doi:10.1016/j.ehb.2003.11.001 (2004).

9 Grasgruber, P., Cacek, J., Kalina, T. & Sebera, M. The role of nutrition and genetics as key determinants of the positive height trend. *Economics & Human Biology* **15**, 81-100, doi:10.1016/j.ehb.2014.07.002 (2014).

10 https://www.ssb.no/a/histstat/tabeller/4-22.html.

11 http://www.dst.dk/pukora/epub/upload/17957/01pop.pdf.

12 K., S. The average height of 18- and 19-year-old conscripts (N = 458,322) in Switzerland from 1992 to 2009, and the secular height trend since 1878. *Swiss Med. Wkly., 141 (2011), p. 13238*.

13 A., S. Body size and lifestyle in an urban population entering adulthood: the 'Grow up Gothenburg' study. *Acta Paediatr., 101 (2012), pp. 964-972*.

14 B., W. Growth in Sweden. *Swedish conscripts, Karolinska Institutet, Stockholm (2007)*.

15 H., K. Height and weight norms for adult Estonian men and women (aged 20–70 years) and ways of somatotyping using a height-weight classification. *Papers on Anthropology, 17 (2008), pp. 113-130*.

16 (NCD-RisC), N. R. F. C. A century of trends in adult human height. .

17 Wang, N. *et al.* Associations of adult height and its components with mortality: a report from cohort studies of 135,000 Chinese women and men. *Int J Epidemiol* **40**, 1715-1726 (2011).

18 Yeboah, J. Diet, height, and health. *Am J Clin Nutr* **106**, 443-444, doi:10.3945/ajcn.117.161562 (2017).

19 Ma, W. *et al.* Adult height, dietary patterns, and healthy aging. *Am J Clin Nutr* **106**, 589-596, doi:10.3945/ajcn.116.147256 (2017).

20 Sawada, N. *et al.* The association between adult attained height and sitting height with mortality in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Plos One* **12**, doi:ARTN e01731170.1371/journal.pone.0173117 (2017).

21 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Diet, Nutrition, Physical Activity and Colorectal Cancer. 2017. Available at: wcrf.org/colorectal-cancer-2017. All CUP reports are available at wcrf.org/cupreports.

22 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Diet, Nutrition, Physical Activity and Breast Cancer. 2017. Available at: wcrf.org/breast-cancer-2017. All CUP reports are available at wcrf.org/cupreports.
23 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Diet, Nutrition, Physical Activity and Kidney Cancer 2014. Available at: wcrf.org/kidney-cancer-2015. All CUP reports are available at wcrf.org/cupreports.
24 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Food, Nutrition, Physical Activity and the Prevention of Pancreatic Cancer 2012. Available at: http://www.dietandcancerreport.org. All CUP reports are available at wcrf.org/cupreports.
25 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Food, Nutrition, Physical Activity and the Prevention of Prostate Cancer 2014. Available at: http://www.dietandcancerreport.org. All CUP reports are available at wcrf.org/sites/default/files/Prostate-Cancer-2014-Report.pdf.
26 Wiseman, M. The Wcrf/Aicr Continuous Update Project: Systematic Review of Mechanistic Studies in Relation to Food, Nutrition, Physical Activity and Cancer. *Annals of Nutrition and Metabolism* **63**, 63-63 (2013).
27 World Cancer Research Fund/American Institute for Cancer Research. Continious Update Project Report: Food, Nutrition, Physical Activity and the prevention of Ovarian Cancer 2014. Available at: http://www.dietandcancerreport.org/cup/cup_resources.php.
28 World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC: AICR, 2007
29 World Cancer Research Fund International.Continuous Update Project. Available at: http://www.wcrf.org/int/research-we-fund/continuous-update-project-cup.
30 Gunnell, D. *et al.* Height, leg length, and cancer risk: a systematic review. *Epidemiol Rev* **23**, 313-342 (2001).
31 Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-838, doi:10.1038/nature09410 (2010).
32 Le Marchand, L. The predominance of the environment over genes in cancer causation: implications for genetic epidemiology. *Cancer Epidemiol Biomarkers Prev* **14**, 1037-1039, doi:10.1158/1055-9965.EPI-04-0816 (2005).
33 Green, J. *et al.* Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. *Lancet Oncol* **12**, 785-794, doi:10.1016/S1470-2045(11)70154-1 (2011).
34 Lin, Z. W. Analysis of the height dependence of site-specific cancer risk in relation to organ mass. *Ann Transl Med* **4**, 88, doi:10.21037/atm.2016.03.04 (2016).
35 Walter, R. B., Brasky, T. M., Buckley, S. A., Potter, J. D. & White, E. Height as an Explanatory Factor for Sex Differences in Human Cancer. *Jnci-J Natl Cancer I* **105**, 860-868, doi:10.1093/jnci/djt102 (2013).

36 Trichopoulos, D. & Lipworth, L. Is Cancer Causation Simpler Than We Thought, but More Intractable. *Epidemiology* **6**, 347-349 (1995).

37 Albanes, D. & Winick, M. Are cell number and cell proliferation risk factors for cancer? *J Natl Cancer Inst* **80**, 772-774 (1988).

38 Okasha, M., Gunnell, D., Holly, J. & Davey Smith, G. Childhood growth and adult cancer. *Best Pract Res Clin Endocrinol Metab* **16**, 225-241, doi:10.1053/beem.2002.0204 (2002).

39 Silventoinen, K. Determinants of variation in adult body height. *J Biosoc Sci* **35**, 263-285 (2003).

40 Gillman, M. W., Gluckman, P. D. & Rosenfeld, R. G. Recent advances in growth research: nutritional, molecular and endocrine perspectives. Preface. *Nestle Nutr Inst Workshop Ser* **71**, IX-X, doi:10.1159/000345196 (2013).

41 Millward, D. J. Nutrition, infection and stunting: the roles of deficiencies of individual nutrients and foods, and of inflammation, as determinants of reduced linear growth of children. *Nutr Res Rev* **30**, 50-72, doi:10.1017/S0954422416000238 (2017).

42 Grasgruber, P., Cacek, J., Kalina, T. & Sebera, M. The role of nutrition and genetics as key determinants of the positive height trend. *Econ Hum Biol* **15**, 81-100, doi:10.1016/j.ehb.2014.07.002 (2014).

43 Hedges, D. W., Berrett, A. N., Erickson, L. D., Brown, B. L. & Gale, S. D. Association between infection burden and adult height. *Econ Hum Biol* **27**, 275-280, doi:10.1016/j.ehb.2017.08.002 (2017).

44 McCrory, C. *et al.* Socioeconomic differences in children's growth trajectories from infancy to early adulthood: evidence from four European countries. *J Epidemiol Community Health* **71**, 981-989, doi:10.1136/jech-2016-208556 (2017).

45 Jeffreys, M., Smith, G. D., Martin, R. M., Frankel, S. & Gunnell, D. Childhood body mass index and later cancer risk: a 50-year follow-up of the Boyd Orr study. *Int J Cancer* **112**, 348-351, doi:10.1002/ijc.20423 (2004).

46 Bjorge, T., Engeland, A., Tverdal, A. & Smith, G. D. Body mass index in adolescence in relation to cause-specific mortality: a follow-up of 230,000 Norwegian adolescents. *Am J Epidemiol* **168**, 30-37, doi:10.1093/aje/kwn096 (2008).

47 Lv, M., Zhu, X., Wang, H., Wang, F. & Guan, W. Roles of caloric restriction, ketogenic diet and intermittent fasting during initiation, progression and metastasis of cancer in animal models: a systematic review and meta-analysis. *PLoS One* **9**, e115147, doi:10.1371/journal.pone.0115147 (2014).

48 Dirx, M. J., Zeegers, M. P., Dagnelie, P. C., van den Bogaard, T. & van den Brandt, P. A. Energy restriction and the risk of spontaneous mammary tumors in mice: a meta-analysis. *Int J Cancer* **106**, 766-770, doi:10.1002/ijc.11277 (2003).

49 Ross, M. H. & Bras, G. Lasting influence of early caloric restriction on prevalence of neoplasms in the rat. *J Natl Cancer Inst* **47**, 1095-1113 (1971).

50 Engelman, R. W., Day, N. K. & Good, R. A. Calorie intake during mammary development influences cancer risk: lasting inhibition of C3H/HeOu mammary tumorigenesis by peripubertal calorie restriction. *Cancer Res* **54**, 5724-5730 (1994).

51 Cleary, M. P. & Grossmann, M. E. The manner in which calories are restricted impacts mammary tumor cancer prevention. *J Carcinog* **10**, 21, doi:10.4103/1477-3163.85181 (2011).

52 Speakman, J. R. & Mitchell, S. E. Caloric restriction. *Mol Aspects Med* **32**, 159-221, doi:10.1016/j.mam.2011.07.001 (2011).

53 Koupil, I. *et al.* Cancer mortality in women and men who survived the siege of Leningrad (1941-1944). *International Journal of Cancer* **124**, 1416-1421, doi:Doi 10.1002/ljc.24093 (2009).

54 Keinan-Boker, L., Vin-Raviv, N., Liphshitz, I., Linn, S. & Barchana, M. Cancer incidence in Israeli Jewish survivors of World War II. *J Natl Cancer Inst* **101**, 1489-1500, doi:10.1093/jnci/djp327 (2009).

55 Vin-Raviv, N., Barchana, M., Linn, S. & Keinan-Boker, L. Severe caloric restriction in young women during World War II and subsequent breast cancer risk. *Int J Clin Pract* **66**, 948-958, doi:10.1111/j.1742-1241.2012.02966.x (2012).

56 Fentiman, I. S., Allen, D. S. & Ellison, G. T. The impact of the Occupation of Guernsey 1940-1945 on breast cancer risk factors and incidence. *Int J Clin Pract* **61**, 937-943, doi:10.1111/j.1742-1241.2007.01288.x (2007).

57 Elias, S. G., Peeters, P. H., Grobbee, D. E. & van Noord, P. A. The 1944-1945 Dutch famine and subsequent overall cancer incidence. *Cancer Epidemiol Biomarkers Prev* **14**, 1981-1985, doi:10.1158/1055-9965.EPI-04-0839 (2005).

58 Elias, S. G., Peeters, P. H. M., Grobbee, D. E. & Noord, P. A. H. v. Breast Cancer Risk After Caloric Restriction During the 1944-1945 Dutch Famine. *JNCI Journal of the National Cancer Institute* **96**, 539-546, doi:10.1093/jnci/djh087 (2004).

59 Da Li, Q. *et al.* Nutrition deficiency increases the risk of stomach cancer mortality. *Bmc Cancer* **12**, doi:Artn 315. Doi 10.1186/1471-2407-12-315 (2012).

60 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Energy restriction early in life and colon carcinoma risk: results of The Netherlands Cohort Study after 7.3 years of follow-up. *Cancer* **97**, 46-55, doi:10.1002/cncr.11052 (2003).

61 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Diet in adolescence and the risk of breast cancer: results of the Netherlands Cohort Study. *Cancer Causes Control* **10**, 189-199 (1999).

62 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Energy restriction in childhood and adolescence and risk of prostate cancer: results from the Netherlands Cohort Study. *Am J Epidemiol* **154**, 530-537 (2001).

63 Heinen, M. M., Verhage, B. A., Goldbohm, R. A., Lumey, L. H. & van den Brandt, P. A. Physical activity, energy restriction, and the risk of pancreatic cancer: a prospective study in the Netherlands. *Am J Clin Nutr* **94**, 1314-1323, doi:10.3945/ajcn.110.007542 (2011).

64 Simons, C. C., van den Brandt, P. A., Stehouwer, C. D., van Engeland, M. & Weijenberg, M. P. Body size, physical activity, early-life energy restriction, and associations with methylated insulin-like growth factor-binding protein genes in colorectal cancer. *Cancer Epidemiol Biomarkers Prev* **23**, 1852-1862, doi:10.1158/1055-9965.EPI-13-1285 (2014).

65 Schouten, L. J., van Dijk, B. A., Lumey, L. H., Goldbohm, R. A. & van den Brandt, P. A. Energy restriction during childhood and early adulthood and ovarian cancer risk. *PLoS One* **6**, e27960, doi:10.1371/journal.pone.0027960 (2011).

66 Hughes, L. A. *et al.* Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. *Int J Epidemiol* **39**, 1333-1344, doi:10.1093/ije/dyq062 (2010).

67 Wu, X. D. *et al.* Combined analysis of genomewide scans for adult height: results from the NHLBI family blood pressure program. *Eur J Hum Genet* **11**, 271-274, doi:10.1038/sj.ejhg.5200952 (2003).

68 Xu, J. F. *et al.* Major recessive gene(s) with considerable residual polygenic effect regulating adult height: Confirmation of genomewide scan results for chromosomes 6, 9, and 12. *American Journal of Human Genetics* **71**, 646-650, doi:Doi 10.1086/342216 (2002).

69 Hirschhorn, J. N. *et al.* Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *American Journal of Human Genetics* **69**, 106-116, doi:Doi 10.1086/321287 (2001).

70 Perola, M. *et al.* Quantitative-trait-locus analysis of body-mass index and of stature, by combined analysis of genome scans of five Finnish study groups. *Am J Hum Genet* **69**, 117-123, doi:10.1086/321286 (2001).

71 Jelenkovic, A. *et al.* Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts. *Sci Rep* **6**, 28496, doi:10.1038/srep28496 (2016).

72 Ku, C. S., Loy, E. Y., Salim, A., Pawitan, Y. & Chia, K. S. The discovery of human genetic variations and their use as disease markers: past, present and future. *J Hum Genet* **55**, 403-415, doi:10.1038/jhg.2010.55 (2010).

73 Frazer, K. A., Murray, S. S., Schork, N. J. & Topol, E. J. Human genetic variation and its contribution to complex traits. *Nat Rev Genet* **10**, 241-251, doi:10.1038/nrg2554 (2009).

74 Soranzo, N. *et al.* Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet* **5**, e1000445, doi:10.1371/journal.pgen.1000445 (2009).

75 Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* **45**, 353-361, 361e351-352, doi:10.1038/ng.2563 (2013).

76 Turnbull, C. *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet* **42**, 504-507, doi:10.1038/ng.586 (2010).

77 Rinella, E. S. *et al.* Genetic variants associated with breast cancer risk for Ashkenazi Jewish women with strong family histories but no identifiable BRCA1/2 mutation. *Hum Genet* **132**, 523-536, doi:10.1007/s00439-013-1269-4 (2013).

78 Jia, W. H. *et al.* Genome-wide association analyses in East Asians identify new susceptibility loci for colorectal cancer. *Nat Genet* **45**, 191-196, doi:10.1038/ng.2505 (2013).

79 Zhang, B. *et al.* Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat Genet* **46**, 533-542, doi:10.1038/ng.2985 (2014).

80 Hirschhorn, J. N. & Lettre, G. Progress in Genome-Wide Association Studies of Human Height. *Hormone Research* **71**, 5-13, doi:Doi 10.1159/000192430 (2009).

81 Weng, L. *et al.* SNP-based pathway enrichment analysis for genome-wide association studies. *BMC Bioinformatics* **12**, 99, doi:10.1186/1471-2105-12-99 (2011).

82 Elkon, R. & Agami, R. Characterization of noncoding regulatory DNA in the human genome. *Nat Biotechnol* **35**, 732-746, doi:10.1038/nbt.3863 (2017).

83 Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S. & Snyder, M. Linking disease associations with regulatory information in the human genome. *Genome Res* **22**, 1748-1759, doi:10.1101/gr.136127.111 (2012).

84 Preuss, C., Riemenschneider, M., Wiedmann, D. & Stoll, M. Evolutionary dynamics of co-segregating gene clusters associated with complex diseases. *PLoS One* **7**, e36205, doi:10.1371/journal.pone.0036205 (2012).

85 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

86 Le Marchand, L. & Wilkens, L. R. Design considerations for genomic association studies: importance of gene-environment interactions. *Cancer Epidemiol Biomarkers Prev* **17**, 263-267, doi:10.1158/1055-9965.EPI-07-0402 (2008).

87 Fedak, K. M., Bernal, A., Capshaw, Z. A. & Gross, S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* **12**, 14, doi:10.1186/s12982-015-0037-4 (2015).

88 Phillips, C. V. & Goodman, K. J. The missed lessons of Sir Austin Bradford Hill. *Epidemiol Perspect Innov* **1**, 3, doi:10.1186/1742-5573-1-3 (2004).

89 Bae, S. *et al.* Causal inference in environmental epidemiology. *Environ Health Toxicol* **32**, e2017015, doi:10.5620/eht.e2017015 (2017).

90 D.S., D. C. S. a. W. Sporadic Breast, Ovarian, or Uterine Cancers as Risk Factors for Colorectal Cancer. *Current Colorectal Cancer Reports 2009, 5:24–29*.

91 Sell, C. Caloric restriction and insulin-like growth factors in aging and cancer. *Horm Metab Res* **35**, 705-711, doi:10.1055/s-2004-814156 (2003).

92 Omodei, D. & Fontana, L. Calorie restriction and prevention of age-associated chronic disease. *FEBS Lett* **585**, 1537-1542, doi:10.1016/j.febslet.2011.03.015 (2011).

93 Roberts, S. B. & Speakman, J. Update on human calorie restriction research. *Adv Nutr* **4**, 563-564, doi:10.3945/an.113.004317 (2013).

94 Nolen, G. A. Effect of various restricted dietary regimens on the growth, health and longevity of albino rats. *J Nutr* **102**, 1477-1493 (1972).

95 Sylvester, P. W., Aylsworth, C. F., Van Vugt, D. A. & Meites, J. Influence of underfeeding during the “critical period” or thereafter on carcinogen-induced mammary tumors in rats. *Cancer Res* **42**, 4943-4947 (1982).

96 Cheney, K. E. *et al.* The effect of dietary restriction of varying duration on survival, tumor patterns, immune function, and body temperature in B10C3F1 female mice. *J Gerontol* **38**, 420-430 (1983).

97 Kritchevsky, D. & Klurfeld, D. M. Caloric effects in experimental mammary tumorigenesis. *Am J Clin Nutr* **45**, 236-242 (1987).

98 Klurfeld, D. M., Welch, C. B., Lloyd, L. M. & Kritchevsky, D. Inhibition of DMBA-induced mammary tumorigenesis by caloric restriction in rats fed high-fat diets. *Int J Cancer* **43**, 922-925 (1989).

99 Klurfeld, D. M., Welch, C. B., Davis, M. J. & Kritchevsky, D. Determination of degree of energy restriction necessary to reduce DMBA-induced mammary tumorigenesis in rats during the promotion phase. *J Nutr* **119**, 286-291 (1989).

100 Kritchevsky, D., Welch, C. B. & Klurfeld, D. M. Response of mammary tumors to caloric restriction for different time periods during the promotion phase. *Nutr Cancer* **12**, 259-269, doi:10.1080/01635588909514025 (1989).

101 Kumar, S. P., Roy, S. J., Tokumo, K. & Reddy, B. S. Effect of different levels of calorie restriction on azoxymethane-induced colon carcinogenesis in male F344 rats. *Cancer Res* **50**, 5761-5766 (1990).

102 Roebuck, B. D., Baumgartner, K. J. & MacMillan, D. L. Caloric restriction and intervention in pancreatic carcinogenesis in the rat. *Cancer Res* **53**, 46-52 (1993).

103 Dogan, S., Rogozina, O. P., Lokshin, A. E., Grande, J. P. & Cleary, M. P. Effects of chronic vs. intermittent calorie restriction on mammary tumor incidence and serum adiponectin and leptin levels in MMTV-TGF-alpha mice at different ages. *Oncol Lett* **1**, 167-176, doi:10.3892/ol_00000031 (2010).

104 Klebanov, S. Can short-term dietary restriction and fasting have a long-term anticarcinogenic effect? *Interdiscip Top Gerontol* **35**, 176-192, doi:10.1159/000096562 (2007).
105 Cleary, M. P. *et al.* Prevention of mammary tumorigenesis by intermittent caloric restriction: does caloric intake during refeeding modulate the response? *Exp Biol Med (Maywood)* **232**, 70-80 (2007).
106 Elias, S. G., Peeters, P. H., Grobbee, D. E. & van Noord, P. A. Transient caloric restriction and cancer risk (The Netherlands). *Cancer Causes Control* **18**, 1-5, doi:10.1007/s10552-006-0080-0 (2007).
107 Mizuno, N. K. *et al.* Combination of intermittent calorie restriction and eicosapentaenoic acid for inhibition of mammary tumors. *Cancer Prev Res (Phila)* **6**, 540-547, doi:10.1158/1940-6207.CAPR-13-0033 (2013).
108 Fair, A. M. & Montgomery, K. Energy balance, physical activity, and cancer risk. *Methods Mol Biol* **472**, 57-88, doi:10.1007/978-1-60327-492-0_3 (2009).
109 Walford, R. L., Harris, S. B. & Gunion, M. W. The calorically restricted low-fat nutrient-dense diet in Biosphere 2 significantly lowers blood glucose, total leukocyte count, cholesterol, and blood pressure in humans. *Proc Natl Acad Sci U S A* **89**, 11533-11537 (1992).
110 Walford, R. L., Mock, D., MacCallum, T. & Laseter, J. L. Physiologic changes in humans subjected to severe, selective calorie restriction for two years in biosphere 2: health, aging, and toxicological perspectives. *Toxicol Sci* **52**, 61-65 (1999).
111 Walford, R. L., Mock, D., Verdery, R. & MacCallum, T. Calorie restriction in biosphere 2: alterations in physiologic, hematologic, hormonal, and biochemical parameters in humans restricted for a 2-year period. *J Gerontol A Biol Sci Med Sci* **57**, B211-224 (2002).
112 Vin-Raviv, N., Dekel, R., Barchana, M., Linn, S. & Keinan-Boker, L. World War II-related post-traumatic stress disorder and breast cancer risk among Israeli women: a case-control study. *Int Psychogeriatr* **26**, 499-508, doi:10.1017/S1041610213002081 (2014).
113 Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Group, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med* **3**, e123-130 (2009).
114 Wells GA, S. B., O'Connell D, Peterson J, Welch V, Losos M. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. (2009).
115 Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software* **36**, 1-48 (2010).
116 Higgins, J. P. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat Med* **21**, 1539-1558, doi:10.1002/sim.1186 (2002).
117 Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557-560, doi:10.1136/bmj.327.7414.557 (2003).
118 Bown, M. J. & Sutton, A. J. Quality control in systematic reviews and meta-analyses. *Eur J Vasc Endovasc Surg* **40**, 669-677, doi:10.1016/j.ejvs.2010.07.011 (2010).
119 Sterne, J. A. C. *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal* **343**, doi:ARTN d4002. 10.1136/bmj.d4002 (2011).
120 Egger, M., Davey Smith, G., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629-634 (1997).
121 Thompson, S. G. & Higgins, J. P. T. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* **21**, 1559-1573, doi:10.1002/sim.1187 (2002).

122 Robsahm, T. E. & Tretli, S. Breast cancer incidence in food- vs non-food-producing areas in Norway: possible beneficial effects of World War II. *Br J Cancer* **86**, 362-366, doi:10.1038/sj.bjc.6600084 (2002).
123 Moller, H. Decreased testicular cancer risk in men born in wartime. *J Natl Cancer Inst* **81**, 1668-1669 (1989).
124 Wanderas, E. H., Tretli, S. & Fossa, S. D. Trends in incidence of testicular cancer in Norway 1955-1992. *Eur J Cancer* **31A**, 2044-2048 (1995).
125 Bergstrom, R. *et al.* Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. *J Natl Cancer Inst* **88**, 727-733 (1996).
126 Tretli, S. & Gaard, M. Lifestyle changes during adolescence and risk of breast cancer: an ecologic study of the effect of World War II in Norway. *Cancer Causes Control* **7**, 507-512 (1996).
127 Thorn, M. *et al.* Trends in colorectal cancer incidence in Sweden 1959-93 by gender, localization, time period, and birth cohort. *Cancer Causes Control* **9**, 145-152 (1998).
128 Becker, N., Muscat, J. E. & Wynder, E. L. Cancer mortality in the United States and Germany. *J Cancer Res Clin Oncol* **127**, 293-300 (2001).
129 Moller, H. Trends in incidence of testicular cancer and prostate cancer in Denmark. *Hum Reprod* **16**, 1007-1011 (2001).
130 Svensson, E. *et al.* Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. *Eur J Cancer Prev* **11**, 489-495 (2002).
131 Richiardi L, B. R., Adami HO, Torrång A, Barlow L, Hakulinen T, Rahu M, Stengrevics A, Storm H, Tretli S, Kurtinaitis J, Tyczynski JE, Akre O. Testicular cancer incidence in eight northern European countries: secular and recent trends. *Cancer Epidemiol Biomarkers Prev.*, 2157-2166 (2004).
132 Svensson, E. *et al.* Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. *Cancer Causes Control* **16**, 215-223, doi:10.1007/s10552-004-3073-x (2005).
133 Jacobsen, R. *et al.* Trends in testicular cancer incidence in the Nordic countries, focusing on the recent decrease in Denmark. *Int J Androl* **29**, 199-204, doi:10.1111/j.1365-2605.2005.00605.x (2006).
134 Borsoi, L., Kunze, U., Kunze, M., Groman, E. & Kundi, M. Trends in mortality and mean age at death from lung cancer in Austria (1975-2007). *Cancer Epidemiol* **35**, 120-125, doi:10.1016/j.canep.2010.06.018 (2011).
135 Chauvenet, M. *et al.* Trends in colorectal cancer incidence: a period and birth-cohort analysis in a well-defined French population. *BMC Cancer* **11**, 282, doi:10.1186/1471-2407-11-282 (2011).
136 Angell-Andersen, E. *et al.* The association between nutritional conditions during World War II and childhood anthropometric variables in the Nordic countries. *Ann Hum Biol* **31**, 342-355, doi:10.1080/03014460410001685304 (2004).
137 Mahabir, S. Association between diet during preadolescence and adolescence and risk for breast cancer during adulthood. *J Adolesc Health* **52**, S30-35, doi:10.1016/j.jadohealth.2012.08.008 (2013).
138 van Noord, P. A. Breast cancer and the brain: a neurodevelopmental hypothesis to explain the opposing effects of caloric deprivation during the Dutch famine of 1944-1945 on breast cancer and its risk factors. *J Nutr* **134**, 3399S-3406S (2004).

139 Land, C. E. *et al.* Incidence of female breast cancer among atomic bomb survivors, Hiroshima and Nagasaki, 1950-1990. *Radiat Res* **160**, 707-717, doi:Doi 10.1667/Rr3082 (2003).
140 Ronckers, C. M., Erdmann, C. A. & Land, C. E. Radiation and breast cancer: a review of current evidence. *Breast Cancer Res* **7**, 21-32, doi:10.1186/bcr970 (2005).
141 Papadopoulos, F. C. *et al.* Age at onset of anorexia nervosa and breast cancer risk. *Eur J Cancer Prev* **18**, 207-211, doi:10.1097/CEJ.0b013e32831bc558 (2009).
142 Hilakivi-Clarke, L. Nutritional modulation of terminal end buds: its relevance to breast cancer prevention. *Curr Cancer Drug Targets* **7**, 465-474 (2007).
143 Brundtland, G. H., Liestol, K. & Walloe, L. Height, weight and menarcheal age of Oslo schoolchildren during the last 60 years. *Ann Hum Biol* **7**, 307-322 (1980).
144 Frankel, S., Gunnell, D. J., Peters, T. J., Maynard, M. & Davey Smith, G. Childhood energy intake and adult mortality from cancer: the Boyd Orr Cohort Study. *BMJ* **316**, 499-504 (1998).
145 Dols M, v. A. A. Voedselvoorziening in Nederland tijdens en onmiddellijk na den Tweeden Wereldoorlog 1940-1945 (Food supply during and after the second World War 1940-1945 in the Netherlands [in Dutch]). *Voeding*, 193-207 (1946).
146 Burger G, S. H., Drummond J. Malnutrition and starvation in Western Netherlands, September 1944-45. Part I and II. *The Hague: General State Printing Office* (1948).
147 Hansen, O. G. Food conditions in Norway during the war, 1939-45. *Proc Nutr Soc* **5**, 263-270 (1947).
148 Strøm, A. *Examination Into the Diet of Norwegian Families During the War-years 1942-1945*. (Exp., 1948).
149 Bennett, M. K. Food and agriculture in the Soviet Union, 1917-48. *J Polit Econ* **57**, 185-198 (1949).
150 Elias, S. G. *et al.* Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. *Int J Cancer* **108**, 628-630, doi:10.1002/ijc.11584 (2004).
151 Hursting, S. D., Smith, S. M., Lashinger, L. M., Harvey, A. E. & Perkins, S. N. Calories and carcinogenesis: lessons learned from 30 years of calorie restriction research. *Carcinogenesis* **31**, 83-89, doi:10.1093/carcin/bgp280 (2010).
152 Anisimov, V. N. & Bartke, A. The key role of growth hormone-insulin-IGF-1 signaling in aging and cancer. *Crit Rev Oncol Hematol* **87**, 201-223, doi:10.1016/j.critrevonc.2013.01.005 (2013).
153 McCarty, C. A. Up-regulation of IGF binding protein-1 as an anticarcinogenic strategy: relevance to caloric restriction, exercise and insulin sensitivty. *Medical Hypothese* (1997).
154 Martin, B. *et al.* Caloric restriction: Impact upon pituitary function and reproduction. *Ageing Res Rev* **7**, 209-224, doi:10.1016/j.arr.2008.01.002 (2008).
155 Thissen, J. P., Ketelslegers, J. M. & Underwood, L. E. Nutritional regulation of the insulin-like growth factors. *Endocr Rev* **15**, 80-101, doi:10.1210/edrv-15-1-80 (1994).
156 Soeters, M. R., Soeters, P. B., Schooneman, M. G., Houten, S. M. & Romijn, J. A. Adaptive reciprocity of lipid and glucose metabolism in human short-term starvation. *Am J Physiol Endocrinol Metab* **303**, E1397-1407, doi:10.1152/ajpendo.00397.2012 (2012).
157 Weaver, I. C. G. *et al.* Epigenetic programming by maternal behavior. *Nat Neurosci* **7**, 847-854, doi:10.1038/nn1276 (2004).

158 Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *P Natl Acad Sci USA* **102**, 10604-10609, doi:10.1073/pnas.0500398102 (2005).

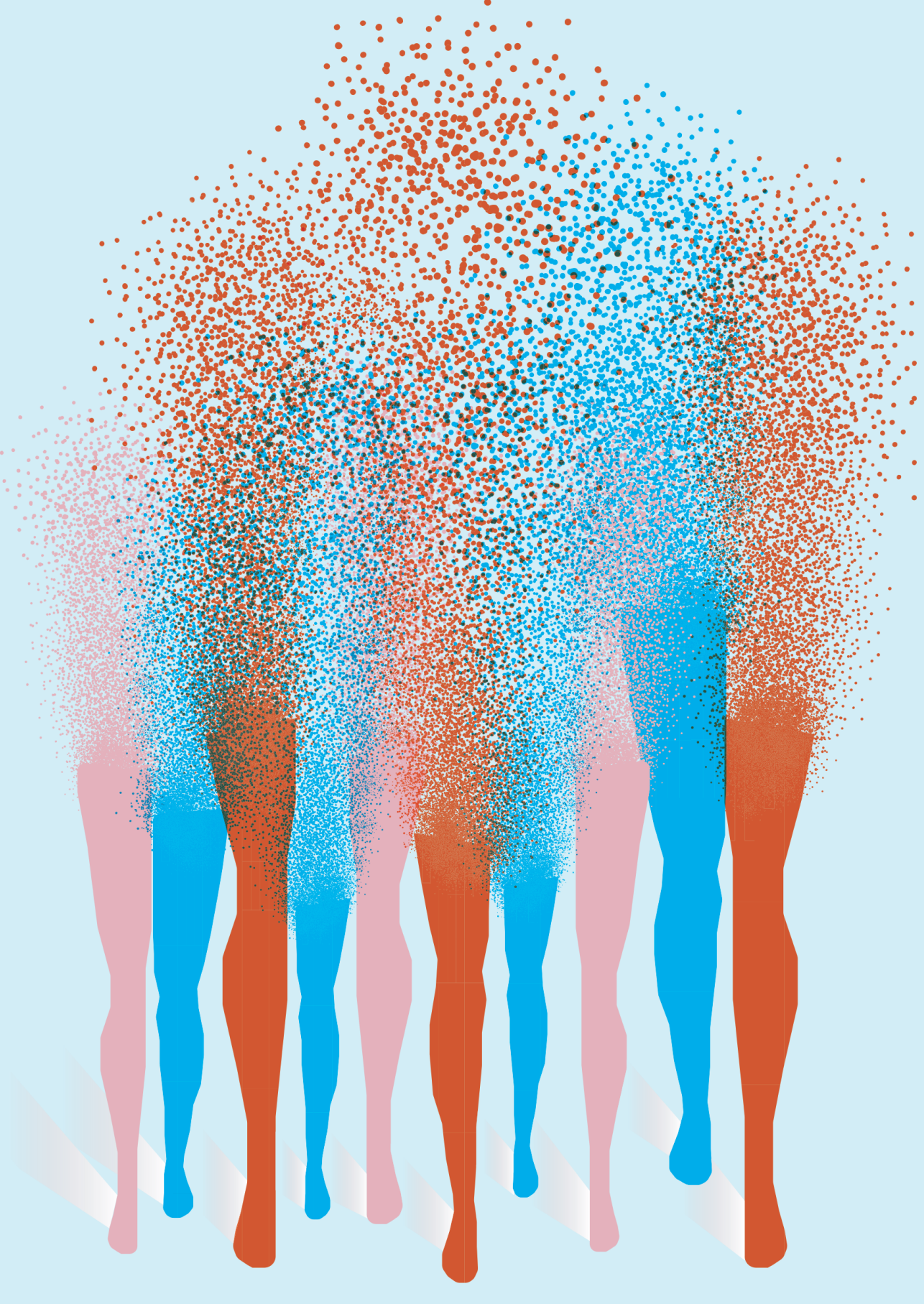
159 De Assis, S. & Hilakivi-Clarke, L. Timing of dietary estrogenic exposures and breast cancer risk. *Ann N Y Acad Sci* **1089**, 14-35, doi:10.1196/annals.1386.039 (2006).

160 Weijenberg, M. P. *et al.* The mTOR Pathway and the Role of Energy Balance Throughout Life in Colorectal Cancer Etiology and Prognosis: Unravelling Mechanisms Through a Multidimensional Molecular Epidemiologic Approach. *Curr Nutr Rep* **2**, 19-26, doi:10.1007/s13668-012-0038-7 (2013).

161 Kim, H. C. *et al.* Aberrant CpG island methylation in early-onset sporadic gastric carcinoma. *J Cancer Res Clin Oncol* **131**, 733-740, doi:10.1007/s00432-005-0017-0 (2005).

162 Kim, H. C. *et al.* CpG island methylation as an early event during adenoma progression in carcinogenesis of sporadic colorectal cancer. *J Gastroenterol Hepatol* **20**, 1920-1926, doi:10.1111/j.1440-1746.2005.03943.x (2005).

163 Hughes, L. A. *et al.* Early life exposure to famine and colorectal cancer risk: a role for epigenetic mechanisms. *PLoS One* **4**, e7951, doi:10.1371/journal.pone.0007951 (2009).



Chapter 2

# A Systematic Literature Review and Meta-regression Analysis on Early Life Energy Restriction and Cancer Risk in Humans

Rachel JJ Elands, Colinda CJM Simons, Martien van Dongen, Leo J Schouten, Bas AJ Verhage, Piet A van den Brandt, Matty P Weijenberg - *PLoS One*. 2016 Sep 19;11(9)

## Abstract

**Background**: In animal models, long-term moderate energy restriction (ER) is reported to decelerate carcinogenesis, whereas the effect of severe ER is inconsistent. The impact of early-life ER on cancer risk has never been reviewed systematically and quantitatively based on observational studies in humans.

**Objective**: We conducted a systematic review of observational studies and a meta-(regression) analysis on cohort studies to clarify the association between early-life ER and organ site-specific cancer risk.

**Methods:** PubMed and EMBASE (1982 – August 2015) were searched for observational studies. Summary relative risks (RRs) were estimated using a random effects model when available ≥ 3 studies.

**Results:** Twenty-four studies were included. Eleven publications, emanating from seven prospective cohort studies and some reporting on multiple cancer endpoints, met the inclusion criteria for quantitative analysis. Women exposed to early-life ER (ranging from 220-1660 kcal/day) had a higher breast cancer risk than those not exposed ($RR_{RE\ all\ ages}$=1.28, 95% CI: 1.05–1.56; $RR_{RE\ for\ 10\text{-}20\ years\ of\ age}$=1.21, 95% CI: 1.09–1.34). Men exposed to early-life ER (ranging from 220 - 800 kcal/day) had a higher prostate cancer risk than those not exposed ($RR_{RE}$=1.16, 95% CI: 1.03–1.30). Summary relative risks were not computed for colorectal cancer, because of heterogeneity, and for stomach-, pancreas-, ovarian-, and respiratory cancer because there were <3 available studies. Longer duration of exposure to ER, after adjustment for severity, was positively associated with overall cancer risk in women ($p$=0.02). Ecological studies suggest that less severe ER is generally associated with a reduced risk of cancer.

**Conclusions:** Early-life transient severe ER seems to be associated with increased cancer risk in the breast (particularly ER exposure at adolescent age) and prostate. The duration, rather than severity of exposure to ER, seems to positively influence relative risk estimates. This result should be interpreted with caution due to the limited number of studies and difficulty in disentangling duration, severity, and geographical setting of exposure.

## Introduction

Energy restriction (ER) without malnutrition has been reported to be the most effective dietary intervention to decelerate aging related diseases [52,91-93], including reductions in cancer risk in animal models of cancer. Lifelong ER starting early in life may be particularly effective in reducing cancer risk at a number of organ sites, predominantly on mammary tumours in rodents [49,50].

Specific aspects of ER, such as the duration and the intensity of ER, may determine whether exposure is associated with an increased or decreased risk for different cancer sites in animal models [51,53]. With regard to the duration of ER, the incidence of neoplasms was reduced following continuous ER throughout lifespan [47-50,94-107], whereas transient ER for several weeks followed by refeeding *ad libitum* has not consistently been associated with the same protective effect and may instead have adverse effects on carcinogenesis [100,104,106]. With regard to the intensity of ER, tumor incidence reduction starts becoming apparent at energy intake below approximately 80% of *ad libitum* levels in spontaneous- [48] and chemically induced tumor models [99]. Several studies have shown the tumor-inhibiting effect of ER to be dose-dependent[97] with the highest protection at about 60% of *ad libitum* energy intake[52,93,108] However, evidence exists for a transition phase of the ER effect: reversal from an increased to a decreased life- and health span [52,93]. Energy intake reduction up to 65% improves life- and health span in rodents, most noticeably by reducing the incidence of multiple forms of cancer, yet it has been suggested that energy intake reduction higher than 65% could not impose the same health benefits [93].

As opposed to the results from controlled animal experimental studies, the scientific evidence for the relationship between ER and cancer risk in humans is inconclusive. Overweight is an established risk factor for many cancers and it is interesting to explore how ER, which is on the other end of the energy balance spectrum, is related to cancer risk, especially given the protective effects of life-long ER in animal models. Short-term experimental studies on voluntarily imposed ER in humans in combination with nutrient dense diets have been conducted to investigate physiological health effects in humans [109-111]. However, investigating long-term effects on cancer risk in human experimental studies is not ethical. Therefore, evidence for associations of ER with cancer in humans is only derived from observational studies. In these studies, ER exposure in humans is mostly early in life and often war-related. This complicates the matter since

extreme conditions may be accompanied by other risk factors; such as stress [112], which may obscure the relationship. In addition, it is obvious that these extreme conditions do not translate directly into prevention, but evidence for such an association points to periods in life that are sensitive to energy balance and its effect on cancer risk decades later.

Existing reviews on human observational research concerning the association between early-life ER and cancer risk have been descriptive in nature. The association between ER in early-life and cancer risk in humans has neither been reviewed systematically nor has it been quantified. This is particularly true for site-specific cancers other than breast cancer. Evidence from *Elias et al., 2005*, who found that overall cancer risk is pulled towards a positive association only when breast cancer cases were included in the analysis [106], further substantiates our objective to study site-specific associations. Therefore, we aimed to review the site-specific associations for ER and cancer risk or -mortality in the literature and, where possible, provide summary relative risk estimates. Comparison of the direction of the site-specific associations will provide insight into whether general or site-specific mechanisms might be involved in human cancer aetiology. Since most studies investigated ER in childhood and adolescence and later life cancer risk, we will focus on this time window. In addition, we aim to investigate in an explorative fashion, how contextual aspects of ER such as timing, duration and severity of early-life ER may impact the reported associations with cancer risk, as has been observed in animal studies.

# Methods

The literature was reviewed for human observational studies on ER in early-life, including adolescence and childhood, in relation to the site-specific cancer risk or mortality in later life until August 2015. PRISMA guidelines for publishing systematic reviews and meta-analysis were followed [113] (S1 Table). The review protocol is described below.

## Search strategy

PubMed and Embase were searched for full-text English-language papers on human observational studies combining the relevant keywords or medical subject headings as follows: '((energy restriction OR famine OR caloric restriction OR World War 2 OR World War II) AND (cancer risk) AND human)/ep)'. References

cited in published original articles were hand-searched until no further studies were identified. Articles were selected only if an abstract was available.

## Study selection

Studies were included in the systematic literature review and meta-analysis if they met the following criteria: 1) study was conducted in a human population; and 2) outcome of interest was site-specific cancer risk or mortality, and effect estimates (hazard ratio (HR), risk ratio (RR) or odds ratio (OR)) with 95% confidence intervals (CIs) were reported or it concerned an ecological study. Studies exclusively on prenatal exposure to ER and ER due to anorexia nervosa were excluded.



## Data abstraction

### Characteristics of included studies

Data were extracted from the included articles by one reviewer (RE). The following information was obtained from the included publications: the first author's last name, publication year, study design, country of origin, cohort size, number of cases, number of person-years of follow-up, age or multivariable adjusted HRs, RRs or ORs and their corresponding 95% CIs, exposure contrasts, estimates of caloric intake, duration of ER, birth cohort, sex, age at exposure and cancer endpoints.

### Methodological quality assessment of included studies

Qualitative assessment of the included cohort studies was examined according to the guidelines in the Newcastle-Ottawa scale (NOS) [114]. The NOS has been typically used for assessing the quality of non-randomized studies in meta-analyses. The NOS contains the following three subscales: selection of the study population (four items), comparability of exposed and non-exposed subcohorts (one item), and outcome assessment (three items). The following characteristics were evaluated: representativeness of the exposed cohort, selection of the non-exposed cohort, ascertainment of the exposure, demonstration that the outcome was not yet present at the start of the study, assessment of the outcome, follow-up time and completeness of follow-up. Quality of included studies was rated by two reviewers (RJJE, CCJMS). A third reviewer (MPW) was counselled in case of any disagreement. The NOS uses a star system to judge studies on key domains. For each domain either a 'star' or 'no star' is assigned, with a 'star' indicating the relevant study design aspect is considered adequate and unlikely to introduce bias. A cohort study can be awarded a maximum of eight stars.

## Meta-analyses

Pooled random effects and 95% CIs were estimated by the restricted maximum-likelihood estimator using the 'metafor' package for R statistical software environment (version 3.1.2) [115]. A random effects model was used, because the cancer (mortality) risk estimates found in the individual studies might be context dependent, due to study-specific characteristics such as duration and severity of ER. Therefore, variation in risk estimates between studies is expected to exceed chance (sampling error) variation, which is accounted for in a random-effects model. We pooled hazard ratios and risk ratios if at least three studies reported on cancer site-specific incidence or mortality and if Higgins' index for between-study heterogeneity ($I^2$) [116] in the reported effect sizes between studies was <50% [117,118]. Heterogeneity was further tested using the Cochran's Q test ($p< 0.1$ indicates statistically significant heterogeneity). In case of statistically significant between-study heterogeneity, we decided to refrain from presenting the pooled relative risk estimate. For these cancer sites we restricted the results presentation to a forest plot visualizing the direction and strength of the associations. In the calculation of pooled effects, the contribution of each study was weighed by the inverse of its variance to take into account study specific variance and variance due to differences in sample size between the studies: $w_i = 1/(v_i + \hat{\tau}^2)$, where $v_i$ denotes the sampling variance (the square root of the standard error) for the given study and $\hat{\tau}^2$ denotes the estimate of (the total amount of heterogeneity between all studies) [115].

If studies were reporting on multiple categories of exposure to early-life ER, the outcomes for the most extreme exposure contrast were included in the meta-analysis. If a cohort reported effect estimates for multiple birth cohorts without an overall estimate, we first pooled estimates of these separate birth cohorts and included the pooled estimate in our meta-analysis. We did so, because the inclusion of multiple effect estimates from the same cohort for a particular endpoint will (artificially) lower the amount of heterogeneity between studies and will drive the pooled estimate into the direction of the findings within one particular cohort, especially in the event of few other cohort studies.

Following recommendations by *Sterne et al.*, 2011, by default, publication bias was evaluated visually only if a minimum of 10 studies were available by inspecting the symmetry of funnel plots [119]. The degree of funnel plot asymmetry was assessed with the Egger's weighted regression test. Absence of publication bias is reflected in an intercept close to 0 with a corresponding $p \geq 0.05$ [120].

Subgroup analyses were conducted where possible for age of exposure to ER. Furthermore, in an explorative fashion, we studied three mixed-effects (meta-regression) models to elucidate whether ER severity and duration, which are inherently linked to the historical setting of the individual included cohort studies, explain part of the variability in effect estimates across studies. We included as explanatory (i.e. independent) variables ER severity, ER duration and ER severity and duration simultaneously, respectively. Since a meta-regression analysis is only advisable in the event of at least 10 individual studies (*i.e.* data points), these analyses were not performed for site-specific cancer outcomes, but all cancer outcomes in men and women respectively [121].

# Results

## Characteristics of included studies

The flow chart of the search strategy is depicted in **Figure 1**. Electronic database search strategy retrieved 228 full-text articles which were all published in English. Fifty-seven review papers were excluded, leaving 171 records to be assessed for eligibility for the systematic review based on title and abstract. Subsequently, 151 records were excluded because the inclusion criteria were not met or because an exclusion criterion was fulfilled, *e.g.* papers exclusively on prenatal ER or anorexia nervosa as reported exposures. One study was excluded [60] because a more recent publication reported on the same association with longer follow-up time [66]. The nineteen remaining records referred to eleven publications on seven cohort studies, seven ecological studies and one case-control study, respectively. Reference-tracking of the nineteen included papers identified five additional ecological studies, resulting in twenty-four full-text articles that met the criteria for full review, some reporting on multiple cancer endpoints. Eight publications emanating from four cohort studies collected data from populations in Europe (two in the Netherlands, one in England and one in Norway) (**Table 1**) [56,58,61-63,66,122]. Three publications emanated from three cohort studies outside Europe: one in China [59], one in Russia [53] and one from Israël [54]. One case-control study was based on a population from Israël [55] (**Table 1**). All twelve ecological studies investigated European populations (**S2 Table**) [123-135].

Specific cancer (mortality) endpoints were reported for breast- [53-56,58,61,122,126], prostate- [53,54,62,129], colorectal- [53,54,66,127,130,132], testicular- [123-125,129,131,133], stomach- [53,59], respiratory-/lung- [53,54,134], pancreas- [63], and ovarian [65] cancer, and multiple

cancer sites [128] (**Table 1** and **S2 Table**). Three studies calculated age standardized rate ratios and 95% confidence intervals by comparing the observed cancer rates in the exposed group with expected cancer rates in the general population, serving as an approximation for the risk in the non-exposed population [54,59,122]. Eight studies calculated hazard ratios [53,56,58,61-63,65,66], one relative risks[54] , and one odd ratios [55].

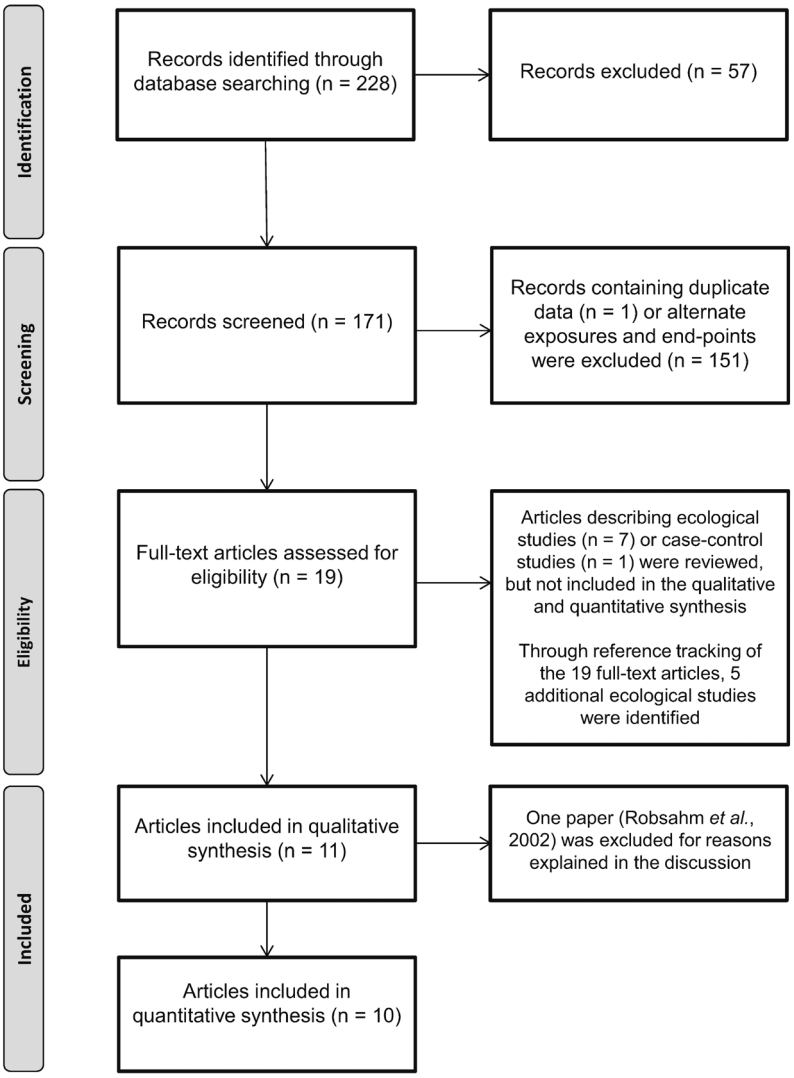



**Figure 1.PRISMA flow diagram showing a breakdown of the study selection.**

**Table 1. Characteristics of included cohort and case-control studies.**

| Studies in chronological order of publication | Design | Country | Historical event | Intervention versus nonintervention cohort/arm (ascertainment of intervention) | Follow-up (years) | Completeness of follow-up | Cases (controls) | N cohort (subcohort) | Adjustments | End-points |
|---|---|---|---|---|---|---|---|---|---|---|
| Dirx et al., 1999 [42] | Case- cohort | The Netherlands | Dutch Hunger Winter | Lived in Western city vs. non-Western area (questionnaire) | 6.3 | >96% | 1009 | 62,573 (1716) | a, b, c, d, e, f | Breast cancer risk |
| Dirx et al., 2001 [43] | Case- cohort | The Netherlands | Dutch Hunger Winter | Lived in Western city vs. non-Western area (questionnaire) | 7.3 | >96% | 903 | 58,279 (1630) | a, d, e, f | Prostate cancer risk |
| Robsahm et al., 2002 [44] | Cohort | Norway | World War 2 | Non-food versus food producing areas (registry) | 28 | - | 7311 | 597,906 | b, d, f | Breast cancer risk |
| Elias et al., 2004 [41] | Case-cohort | The Netherlands | Dutch Hunger Winter | Hunger vs. no hunger (questionnaire) | 15.3 (median) | 95% | 585 | 15,396 (2352) | a, b, d, e | Breast cancer risk |
| Fentiman et al., 2 007 [45] | Cohort | England | Occupation of Guernsey | Stayed vs. evacuated (questionarre) | 15-20 | - | 97 | 2,377 | a, b | Breast cancer risk |
| Koupil et al., 2009 [7] | Cohort | Russia | Siege of Leningrad | Lived in Leningrad vs. outside Leningrad (registry) | 23-30 | ~95% | 792 | 5330 | a, b, c, d, e | Breast-, prostate-, stomach-, colorectal-, respiratory-, other cancers and all-site cancer mortality |
| Keinan-Boker et al., 2009 [48] | Cohort | Israël | Holocaust | Immigrated after the war vs. before the war (registry) | 21 | ~93% | 69,297 | 315,544 | - | Breast-, prostate-, stomach-, colorectal-, lung and bronchial-, other cancers and all-site cancer risk |
| Hughes et al., 2010 [63] | Case- cohort | The Netherlands | Dutch Hunger Winter | Lived in Western city vs. non-Western area (questionnaire) | 16.3 | >96% | 2971 | 120,852 (3981) | a, c, d, e, f | Colorectal cancer risk |
| Heinen et al., 2011 [46] | Case- cohort | The Netherlands | Dutch Hunger Winter | Lived in Western city vs. non-Western area (questionnaire) | 13.3 | >96% | 446 | 120,852 (4774) | a, b, c, d, e, f | Pancreatic cancer risk |
| Schouten et al., 2011 [64] | Case- cohort | The Netherlands | Dutch Hunger Winter | Lived in Western city vs. non-Western area (questionnaire) | 16.3 | >96% | 394 | 62,573 (2589) | a, b | Ovarian cancer risk |
| Li et al., 2012 [47] | Cohort | China | Chinese Famine during Great Leap Forward | Born between 1930-1964 vs. born between 1965-1999 (registry) | 4 | - | 162 | - | - | Stomach cancer mortality |
| Vin-Raviv et al., 2012 [49] | Case-control | Israël | Holocaust | Hunger vs. no hunger (structured interview) | - | - | 65 (200) | - | d | Breast cancer risk |

a Anthropometric variables (body mass index and height),
b reproductive variables (parity, age at first birth, age at menopause, hormone replacement therapy (never, ever), oral contraceptive use (ever, never), hysterectomy (yes, no)),
c smoking or alcohol consumption,
d socio-economic variables (education and economic status),
e genetic factors (family history or genetic mutation tests) or
f variables that indicate baseline energy consumption or energy expenditure (physical activity and energy intake).

## Exposure to energy restriction

All of the included prospective studies investigated exposures to war-related ER except for the Chinese study (**Table 1**) [59]. In most cohorts, exposure to ER was proxy-assessed using information on residential history during the war years from self-reports or registries [53,54,56,61-63,65,66,122]. In one study, exposure measurement was based on residential status and individual recall of severity of exposure to wartime ER [58]. A case-control study used interviewing techniques [55].

The estimated level of caloric intake was retrieved from historical references included in the prospective studies' reports, and ranged from 220 kcal/day [54] to 1660 kcal/day [56] (**S3 Table**). With regard to ER severity, these historical references indicated either states of malnutrition [53,54,59,61-63,65,66] or (semi-)malnutrition [56] during early-life ER. One cohort study from Norway reported moderate early-life ER with a nutritious balanced diet [122]. The duration of exposure to ER ranged from 5-6 months in the Netherlands Cohort Study on Diet and Cancer (NLCS) [58,61-63,65,66] to 72 months in the Jewish Cohort Study [54] (S3 Table).

## Methodological quality assessment of included cohort studies

Methodological quality assessment according to the NOS indicated that the total number of points assigned to each cohort study ranged between 6-7 on a 0-8 scale (**S4 Table**). Most studies failed to receive a point for the item 'ascertainment of exposure", which relates to the fact that most studies had to rely on proxy-assessment of war-related ER. Sensitivity analyses concerning the quality of the included studies were not conducted since the studies were comparable and of high quality.

## Association between early-life ER and site-specific cancer risk

Three or more studies on ER and site-specific cancer risk were available for breast cancer (**Figure 2**), prostate cancer (**Figure 3**), and colorectal cancer in men and women (**Figure 4**), but not stomach, pancreatic and respiratory cancers in men and women, and ovarian cancer in women. For all sites, information on the risk ratios and hazard ratios extracted from the reports is provided in **S5, S6** and **S7 Tables**.

## Breast cancer

All but one of the five prospective cohort studies on early-life ER and breast cancer risk reported an association with increased risk of breast cancer although only significant in two studies (**Figure 2** and **S5 Table**) [53,54,56,58,61]. Pooling the

risk estimates for breast cancer from these five prospective cohort studies on ER between in utero - 33 years of age showed a significantly increased risk ($RR_{RE}$=1.28, 95% CI: 1.05-1.56, $I^2$=49.89%; *p*= 0.08 for Cochran's Q test) (**Figure 2**). ER exposure was between 220-1660 kcal/day. A meta-analysis could also be conducted for ER exposure between 10 and 20 years of age as shown in **Figure 2**. Women exposed to ER between 10 and 20 years of age had significantly increased risk of breast cancer compared to those not exposed during that age period ($RR_{RE}$=1.21, 95% CI: 1.09–1.34, $I^2$ = 0% for Cochran's Q, *p*= 0.68). We refrained from pooling relative risk estimates for ER exposure between 0-10 years of age, because the Cochran's Q test indicated statistically significant between-study heterogeneity ($I^2$=63.28%; *p*= 0.05 for Cochran's Q) (**Figure 2**). The study by *Robsahm et al.*,[122] could not be included in the meta-analysis since for the exposure contrast that was investigated, i.e. non-food versus food producing areas, there were already differences in absolute cancer incidence that already existed before war-related exposure occurred. The findings supported an increased breast cancer risk among birth cohorts that were of adolescent age and living in non-food producing areas during WWII, who were exposed to ER compared to food-producing areas. One case-control study on early-life ER during the Holocaust and breast cancer risk reported an increased risk for ER exposed women [55]. In contrast, two ecological studies reported findings suggesting an inverse association for ER with breast cancer incidence [126] and mortality (**S2 Table**) [128]. A drop in breast cancer incidence rates was observed in Norwegian women exposed to war-time related ER (intake approx. 20% restricted [136]) during puberty [126]. Similarly, breast cancer mortality was low in women in early post-war Germany but increased afterwards comparable to levels in the United States. These women born around the war years in Germany were restricted to an estimated 1412-1600 kcal/day intake in 1945 when food supplies were plummeting (**S2 Table**) [128].

## Prostate cancer

Three prospective cohort studies on early-life ER and prostate cancer risk and mortality indicated that men exposed to ER have a higher prostate cancer risk compared to those not exposed [53,54,62], although only significant for one study. Results from the meta-analysis indicate that men exposed to ER (energy intake estimates ranging from 220 - 800 kcal/day) had a significantly increased prostate cancer risk compared to non-exposed men ($RR_{RE}$ = 1.16, 95% CI: 1.03–1.30; $I^2$=0%; *p*= 0.84 for Cochran's Q) (**Figure 3** and **S6 Table)**.

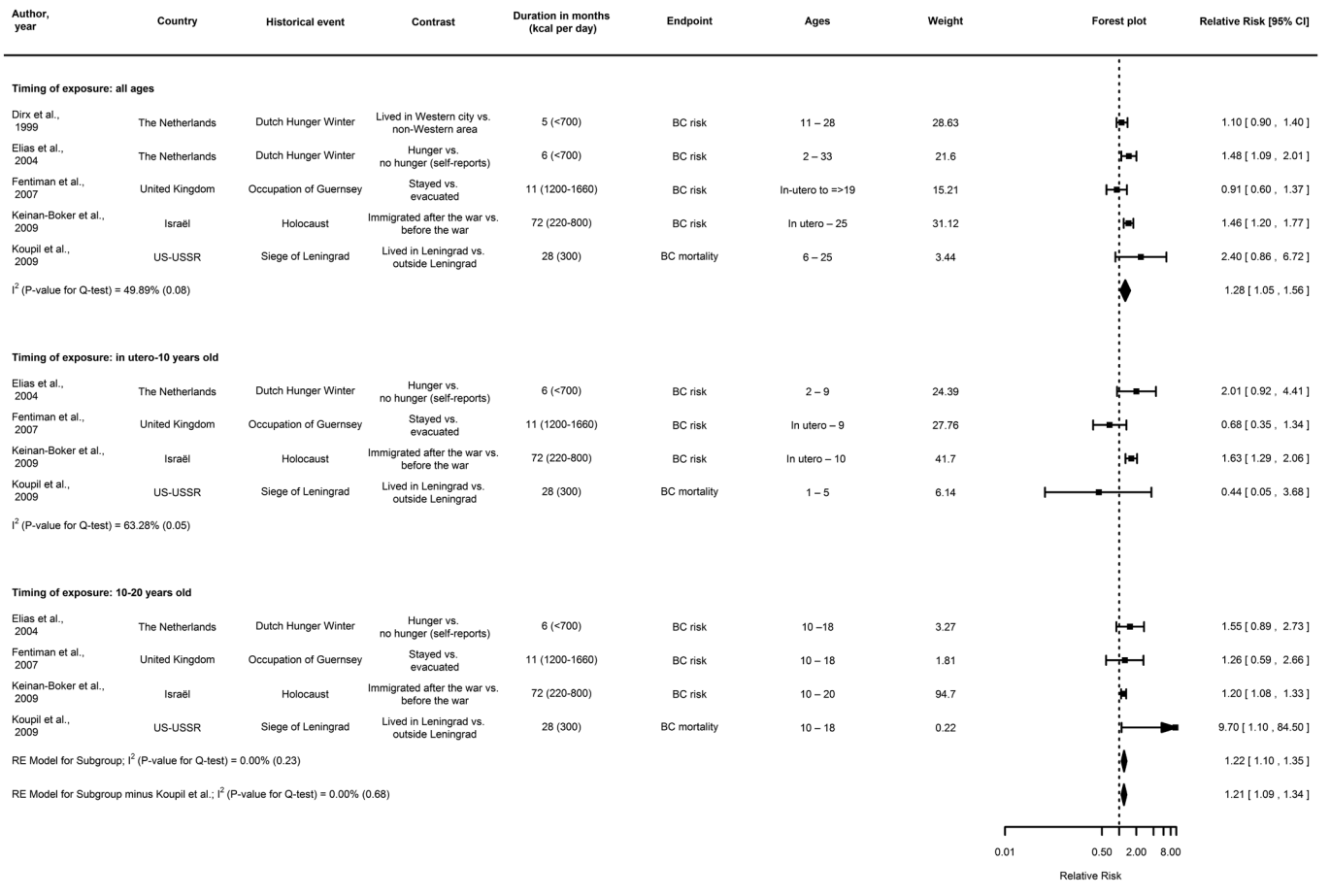



**Figure 2. Forest plot showing a meta-analysis of cohorts on the association between transient early-life energy restriction and the relative risk and mortality of breast cancer, using the relative risk estimate as summary statistic**. Note: Subgroup analyses were performed for childhood (in utero-10 years old) and adolescent (10–20 years old) exposure to ER in relation to breast cancer risk. If individual studies provided risk ratio estimates for different birth cohorts, these were pooled and the pooled estimate was taken along in the meta-analysis. Abbreviations: CI, confidence interval;BC, breast cancer.

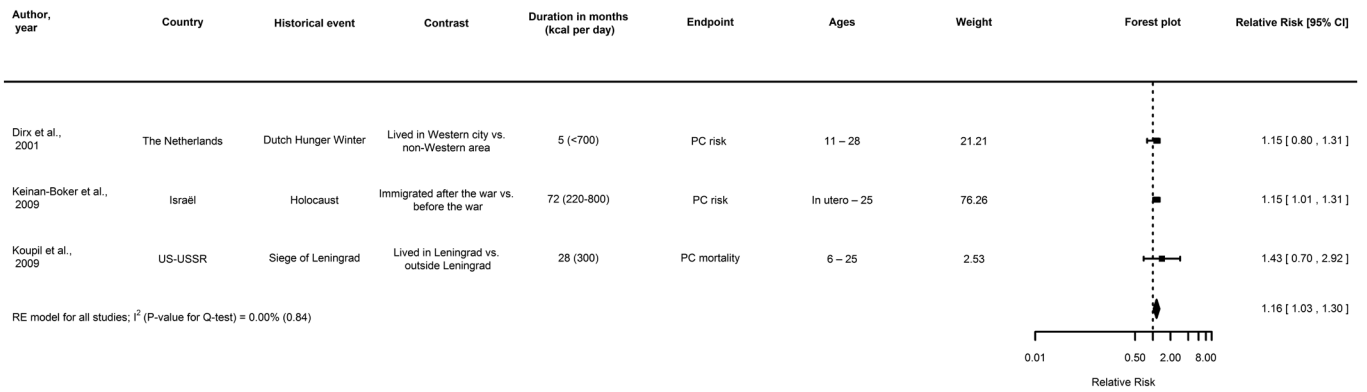



**Figure 3. Forest plot showing a meta-analysis of cohorts on the association between transient energy restriction during (pre) adolescence and the relative risk and mortality of prostate cancer.** Note: If individual studies provided risk ratio estimates for different birth cohorts, these were pooled and the pooled estimate was taken along in the meta-analysis. Abbreviations: CI, confidence interval;PC, prostate cancer.

In contrast, two ecological studies reported findings suggesting an inverse association between ER and prostate cancer incidence [129] and mortality [128] (**S2 Table**). One ecological study was conducted in Denmark and reported a low point in prostate cancer incidence after the Second World War [129] in individuals potentially subjected to an estimated 7% reduction in energy intake [136] (**S2 Table**). Similarly, prostate cancer mortality was low in males in early post-war Germany, but increased afterwards comparable to levels in the United States; men were subjected to an estimated 1412-1600 kcal/day in 1945 compared to those borne earlier or later [128] (**S2 Table**).

## Colorectal cancer

Three cohort studies on colorectal cancer reported positive (in men and women) [54], inverse (in men only) [66] and null associations [53] with early-life ER (**Figure 4** and **S7 Table**). We refrained from pooling the risk estimates for colorectal cancer from these three prospective cohort studies on ER, due to statistically significant between-study heterogeneity in men ($I^2$=90.02%; $p<$ 0.001 for Cochran's Q test), (**Figure 4**) and women ($I^2$=87.96%; $p<$ 0.001 for Cochran's Q test), (**Figure 4**). The study on childhood and adolescent ER during the Holocaust reported associations with increased colorectal cancer risk in both men and women [54]. One prospective cohort study was on adolescent ER during the Dutch Hunger Winter and its association with proximal, rectal and overall colorectal cancer incidence demonstrating an association with decreased colorectal cancer risk in men, but no association in women (**Figure 4** and **S6** and **S7 Tables**) [66]. In the study on childhood and adolescent ER during the siege of Leningrad, a non-significant decreased colorectal cancer risk was observed in both men and women [53].

Two ecological studies indicated a drop in age-standardized incidence for colorectal cancer in birth cohorts encompassing the period of the Second World War in Norway, Sweden, Denmark and Estonia, but this drop in estimated colorectal cancer incidence did not extend to Finland (**S2 Table**) [130,132]. The drop in absolute colorectal cancer incidence in Norway was most pronounced for localizations in the proximal colon for the birth cohorts 1939-1948 for men and 1944-1953 for women [130]. Also, men and women born in Norway between 1944 and 1948 seemed to have a lower risk for cancer of the distal colon and rectum than was expected on the basis of the general trend [130]. An ecological study conducted in Sweden, reported that the relative risk of right-sided colon cancer leveled off in men and women born after 1930, whereas left-sided colon cancer incidence was constant in cohorts born until 1930 and decreased later (**S2 Table**) [127].

## Stomach cancer

There are two prospective studies on early-life ER and stomach cancer mortality (**Figure 4** and **S7 Table**). The first study on childhood and adolescent ER and stomach cancer mortality during the siege of Leningrad reported null associations for both men and women [53]. The second study on childhood ER during the Chinese economic depression and stomach cancer mortality observed a positive association in both men and women [59].

## Pancreatic cancer

One cohort study on adolescent exposure to ER during the Dutch Hunger Winter and pancreatic cancer risk reported no associations in men and women (**Figure 4**) [63].

## Lung cancer

There are two cohort studies reporting on early-life ER and lung cancer risk (**Figure 4** and **S7 Table**). The study on childhood and adolescent ER during the Holocaust and lung cancer risk showed associations with increased lung cancer risk in both men and women[54] The study on childhood and adolescent ER and lung cancer mortality during the siege of Leningrad showed null associations in both men and women [53]. One ecological study showed an increased lung cancer risk in men and women born during or after the Second World War in Austria; overall, there was a decreasing risk in men, but not women, with increasing birth year [134]. However, it is difficult to disentangle changes in smoking habits from other exposures, *e.g.* starvation [134].

## Testicular cancer

Age-period-cohort analyses in ecological studies have indicated reduced testicular cancer incidence rates, interrupting a trend of increasing incidences over time, for cohorts born during the Second World War in Norway, Sweden, and Denmark, but not in Finland (**S2 Table**) [123-125,129,131,133].

## Ovarian cancer

One cohort study on adolescent exposure to ER during the Dutch Hunger Winter and ovarian cancer risk showed no association in women (**Figure 4** and **S7 Table**) [65].

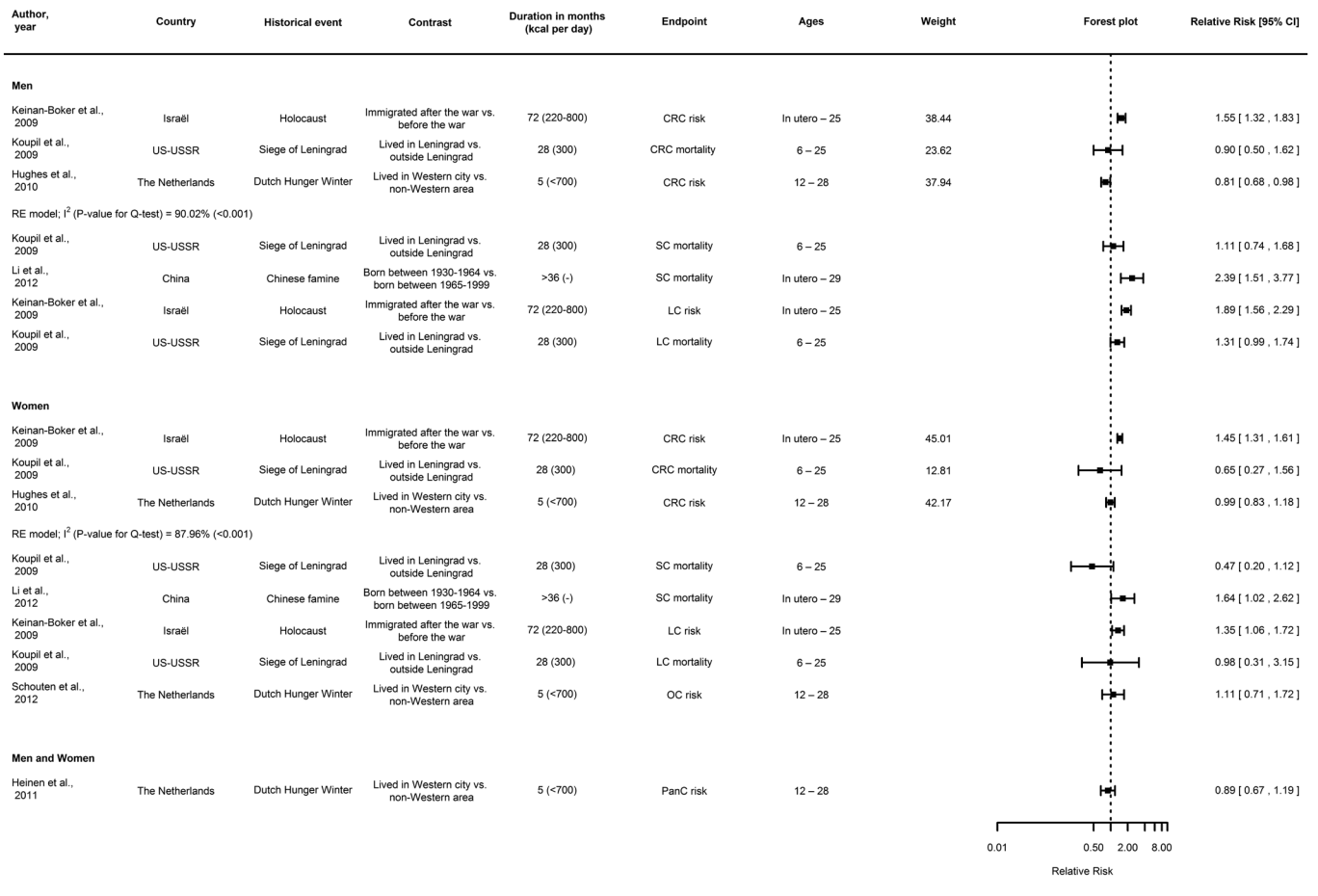



**Figure 4. Forest plot showing a meta-analysis of cohorts on the association between transient energy restriction during (pre) adolescence and the relative risk and mortality at sites other than prostate cancer risk in males and breast cancer risk in females.**

Note: If individual studies provided risk ratio estimates for different birth cohorts, these were pooled and the pooled estimate was taken along in the meta-analysis. Abbreviations: CI, confidence interval; CRC, colorectal cancer; SC, stomach cancer; PaC, pancreatic cancer; LC, lung cancer ;OC, ovarian cancer.

## Duration, severity and timing of ER

The contextual aspects of ER such as duration and severity of early-life ER are an inherent characteristic of the individual studies and these contextual aspects may impact the reported associations between early-life ER and cancer risk. Due to the limited number of studies available it was not possible to disentangle these effects for the different cancer sites separately. To estimate whether between-study heterogeneity was explained by the covariates duration of ER and severity of ER a mixed-effects meta-regression model was fitted across all cancer sites for men and women. A longer duration of exposure to early-life ER (in months) was associated with a (borderline) increased overall cancer risk in men ($p$= 0.07) and women ($p$< 0.001) (**Table 2** and **Figure 5**). The associations were statistically significant after adjusting for severity of exposure in women ($p$< 0.001) but not in men ($p$= 0.08) (**Table 2**). Particularly, in women, adding duration of ER to the model substantially reduced heterogeneity between cohort studies in the meta-analysis (**Table 2**). Severity of ER was not associated with the reported effect size in cohort studies in men ($p$= 0.54) and women ($p$= 0.20) (**Table 2** and **Figure 5**). Yet, overall cancer risk in women tended to increase as the caloric intake per day decreased (**Table 2** and **Figure 5**).

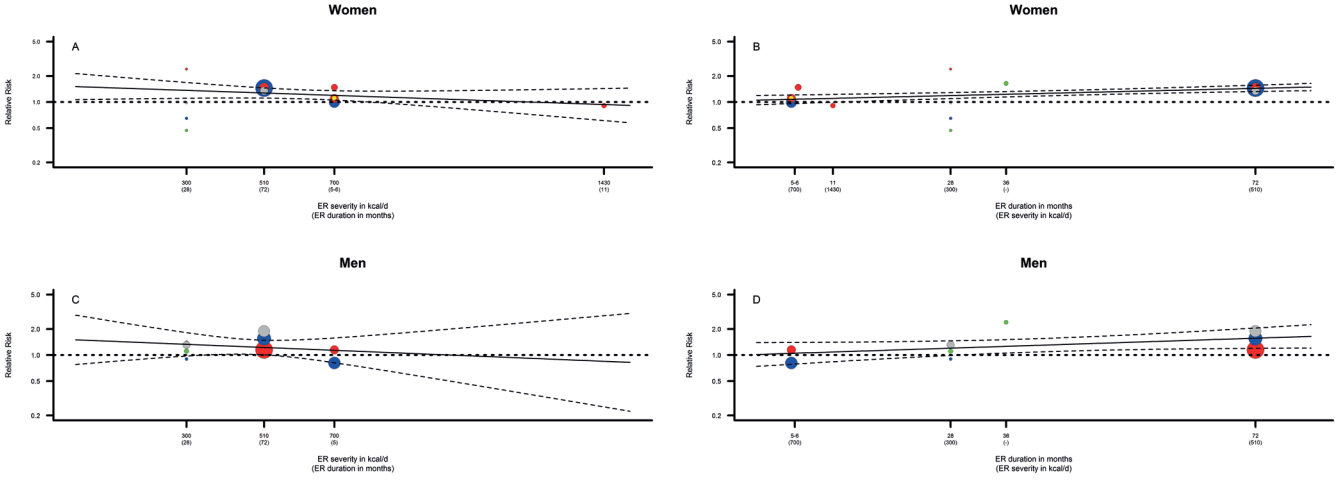



**Figure 5. An overview of some of the contextual aspects of energy restriction that might modulate the association of early-life energy restriction with cancer risk.** Note: The estimated caloric intake (in units of 100 kcal/day) was based on the mid-point caloric intake reported in the publications and was plotted against the reported relative risk ratios from the individual studies separately for women (panel A) and men (panel C). The estimated duration of ER (in months) was plotted against the reported relative risk ratios from the individual studies separately for women (panel B) and men (panel D). In women, the data points indicated in red represent studies reporting on breast cancer risk or mortality; the data points indicated in blue represent studies reporting on colorectal cancer risk or mortality; the data points indicated in green represent studies reporting on stomach cancer risk or mortality; the data points indicated in grey represent studies reporting on lung cancer risk or mortality; and the data points indicated in yellow represent a report on ovarian cancer risk. In men, the data points indicated in red represent studies reporting on prostate cancer risk or mortality; the data points indicated in blue represent studies reporting on colorectal cancer risk or mortality; the data points indicated in green represent studies reporting on stomach cancer risk or mortality; and the data points indicated in grey represent studies reporting on lung cancer risk or mortality. The dashed lines indicate the confidence intervals of the meta-regression line.

**Table 2. Meta-regression for exposure to early-life energy restriction and all type cancer risk/mortality including moderators.**

| Endpoint | Mixed-effects model unless otherwise specified | Beta intercept | 95% CI | Beta | 95% CI | $I^2$ | $R^2$ | Test for heterogeneity | Test for residual heterogeneity | Test of moderators |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | *p-value* | *p-value* | *p-value* |
| All cancers women | RE model | | | 0.20 | (0.07, 0.34) | 57.41% | | 0.002 | | |
| | RR ~ severity of exposure | 0.41 | (0.06, 0.76) | -0.03 | (-0.08, 0.02) | 48.06% | 37.40% | | 0.01 | 0.20 |
| | RR ~ duration of exposure | 0.05 | (-0.07, 0.17) | <0.01 | (0, 0.01) | 0.02% | 99.99% | | 0.15 | <0.001 |
| | RR ~ severity of exposure + | 0.10 | (-0.31, 0.51) | -0.01 | (-0.06, 0.04) | 0.00% | 100.00% | | 0.11 | <0.001 |
| | duration of exposure | | | <0.01 | (0, <0.01) | | | | | |
| All cancers men | RE model | | | 0.26 | (0.06, 0.46) | 84.42% | | <0.001 | | |
| | RR ~ severity of exposure | - | | - | | 83.01% | 0.00% | | <0.001 | 0.54 |
| | RR ~ duration of exposure | -0.01 | (-0.30, 0.33) | 0.01 | (>-0.01, 0.01) | 77.73% | 29.26% | | <0.001 | 0.07 |
| | RR ~ severity of exposure + | 0.05 | (-0.63, 0.73) | <0.01 | (>-0.01, <0.01) | 73.92% | 36.64% | | <0.001 | 0.08 |
| | duration of exposure | | | 0.01 | (0, 0.01) | | | | | |

Note: The unit increases in severity of exposure and duration of exposure were 100 kilocalories per day and months, respectively; severity and duration of exposure were inversely correlated in women and men (r = -0.36 and -0.12, respectively; p = 0.24 and 0.76, respectively). Abbreviations: CI, confidence interval; RE model, random-effects model; RR, relative risk. * Estimates are not shown, because the test of moderators was not statistically significant

# Discussion

The epidemiological evidence for a sustained effect of transient (pre)adolescent ER on site-specific cancer risk has been inconclusive and not been reviewed or quantified previously. In this systematic review and meta-analysis of observational studies, the pooled results of cohort studies indicate that women exposed to ER (energy intake ranging from 220 - 1660 kcal/day) during childhood and adolescence have a 28% increased breast cancer risk. Also, pooled results from cohort studies indicate that exposure to ER (energy intake ranging from 220 - 800 kcal/day) during childhood and adolescence is associated with a 16% increased risk of prostate cancer. Summary risk estimates for colorectal-, stomach-, pancreatic-, ovarian- and respiratory cancer could not be calculated due to the limited number of studies available or study heterogeneity. Meta-regression analyses were conducted across all cancer sites and suggested that a longer duration of exposure (in months) to early-life ER is (borderline) associated with increased cancer risk in women and men. Particularly, in women, between-study heterogeneity was explained by the duration of early-life ER. The associations remained statistically significant in women after adjusting for severity of exposure. Of note is that the results from the meta-regression analysis are exploratory and should be interpreted with caution given that for women only 6 cohorts with 13 risk estimates were included, and for men only 4 cohorts with 10 risk estimates, resulting in a limited power to discriminate between different covariates. The meta-regression analysis showed that the effect sizes in women tended to increase with a decrease of daily caloric intake. This trend was not significant, however. The lack of cohort studies that have investigated more moderate exposures to early-life ER may have obscured a possible relation.

## Inconsistencies between human observational studies

The most obvious finding emerging from this review is the inconsistency of the observed associations between early-life ER and site-specific cancer incidence obtained from various types of human observational studies. Ecological studies suggest either no effect or decreased site-specific cancer risk after transient exposure to severe early-life ER, whereas, prospective cohort studies suggest no effects or increased site-specific cancer risk. There are several potential reasons for the discrepancies between observational studies such as the unique historical contexts and residual confounding from baseline geographical differences in cancer incidence and from other exposures related to war-related uncontrolled ER.

The unique historical settings of the observational studies are associated with geographic location and with the duration and severity of ER. Certain aspects of ER, *i.e.* the timing of exposure [61,62,137,138], its duration and/or severity [54,55,106], may determine whether ER is associated with an increased or decreased risk for different cancer sites. Animal studies have indicated that continuous ER may be particularly effective in reducing cancer risk when started early in life [52]. Our meta-analysis indicated that women exposed to severe transient ER between 10 - 20 years of age were at increased risk of breast cancer, whereas no consistent associations were observed for women exposed between 0 - 10 years of age. Particularly adolescence has been suggested to coincide with a period in which the developing mammary gland is sensitive to environmental signals [139,140]; this has also been observed for exposure to nutritional stimuli, for example, transient severe ER [61,137,141,142]. Regarding the duration of early-life ER, evidence from animal studies indicated that transient ER followed by refeeding *ad libitum* may have adverse effects on carcinogenesis[97,104,106] as opposed to continuous ER [47,48]. Most human studies investigated transient exposures to early-life ER; and in some studies, but not all, reduced food intake persisted for several years after ER exposure [53-55,59] . Also there is evidence concerning the severity of ER; a transition phase of ER may exist between 40% to 65% of daily regular caloric intake, at which the effect of ER reverses from an increase to a decrease of life and health span [52,93]. Typically, the exposures to early-life ER in prospective studies were severe (energy intake estimates ranging from 220 - 1660 kcal/day, corresponding with a reduction in daily energy intake compared to current common daily allowances of 2,000 kcal in adults ranging from 17 - 89%) and coincided with severe ER (>40%) in all but one of the studies [56]. In contrast, ecological studies investigated exposures to moderate ER that were mainly experienced in Denmark, Norway, Sweden and Finland where populations were exposed to an estimated reduction of 4-20% or less of daily caloric intake [136], accompanied by a nutritionally balanced diet [123-127,129-133]. Even though in ecological studies individual data on exposure of the cancer cases are lacking, it can be assumed that the observed reductions in anthropometric measures, *e.g.* weight and height, during the WW-II years in Europe are approximately reflecting the prevailing nutritional conditions in those countries [136,143]. The inverse associations between ER and cancer risk found in ecological studies suggest that moderate ER with adequate nutritional balance could exert a protective effect on cancer whereas more extreme exposure to ER, as reported in prospective cohort studies, might convey a higher cancer risk. This suggestion is supported by a study examining the effects of long-term moderate caloric intake reduction in children and adolescents in Pre-War

Britain that resembles the evidence from human ecological studies and animal experimental models that continuous moderate ER may exert a protective effect on cancer mortality [144].

Another potential reason for the difference in findings between observational studies is that many prospective studies do no account for existing baseline differences in absolute cancer incidence across exposure groups. In prospective studies, often a geographical contrast within a country, *e.g.* food-producing 'rural' areas versus non-food producing 'urban' areas, was employed as a proxy for unrestricted vs. restricted energy intake [53,58,60-63,65,66,122]. These geographical contrasts may include longstanding differences in absolute cancer incidence that existed already before the war-related exposure occurred. For example, *Robsahm et al.* [122] observed a higher cancer incidence in urban areas as compared to rural areas. These geographical differences in cancer incidence may partly result from the different distribution of cancer related risk factors. Since, ecological studies applied temporal contrasts inferred from age-period-cohort modelling these studies were not impacted by geographical differences in absolute cancer incidence; this might explain in part the contrasting findings from ecological and prospective cohort studies. Longstanding baseline differences in cancer risk between geographical areas (*i.e.* urban and rural areas) often coincide with the groups that are contrasted in terms of ER. This may mask a true effect of ER on outcome and may thus have caused attenuation of any true inverse associations that may now remain unobserved or even be reversed revealing positive associations. This potential bias may have resulted in the observation that ER is accompanied with an increased risk of breast and prostate cancer in the meta-analysis. Therefore, caution is warranted in interpreting the results from observational epidemiologic studies on early-life ER in relation to cancer.

Furthermore, exposure to war-related ER is potentially accompanied with other risk factors for cancer, such as stress, which may explain the observed positive associations between more severe early-life ER and cancer risk, and thereby contribute to the difference in findings between cohort and ecological studies. For example, it has been reported that post-traumatic stress disorder in exposed Jewish children during Holocaust suffering from severe ER was associated with increased breast cancer risk [112].

## Mechanistic evidence

Some findings from the limited number of animal studies that have investigated the cancer-related effects of transient severe ER early in life followed by *ad libitum* food consumption are supportive[97,104,106] of the null and positive findings from human prospective studies. Still, while animal experimental studies find inverse associations, in some cases, such as the Dutch Hunger winter, the counteracting increased caloric intake following the famine, might have obscured associations. Whereas the food availability after the war recovered quickly in the Netherlands [145,146] and Norway [147,148], constraints in food availability sustained during the post-war period in the Soviet Union [149]. It has been argued that transient severe ER followed by acute access to abundant food imposes an overshoot of mitogenic growth hormone factor signaling [150], through the growth hormone-insulin-like growth factor (GH-IGF) axis and may result in a modest acceleration of the carcinogenic response in animals [105] and humans [150]. In contrast, continuous moderate ER enables the body's metabolism to adapt on the long-term by responding with lower circulating IGF-1 [52,151,152] and upregulation of IGF binding protein (IGFBP)-1 levels [153] which may suppress carcinogenesis. In general, together with the hypothalamic-pituitary-gonadal axis [138,154], the GH-IGF axis coordinates growth and development early in life, a time during which serum levels of these hormones peak under *ad libitum* conditions [38]. When ER occurs early in life, a period in which development and appropriate functioning of the reproductive axis demands a fixed quantity of energy stores [154], these axes might be permanently modified, influencing cancer risk later in life. Yet, for the GH-IGF-1 axis it is known that the response to ER is different between species. Whereas in both rodents and humans, serum IGF-1 levels decrease [155] and result in a concomitant reduction in growth hormone (GH) serum levels in rodents, GH serum levels tend to increase in humans [153,156]. The contrasting fasting response between species may lead to differences in the observed associations between early-life ER and cancer risk in humans and animal models of carcinogenesis. Correspondingly, an experimental study in humans with a two-year caloric intake restriction of 30% from ad libitum, which resembled the controlled setting of moderate ER with nutrient dense diets in animal experiments, observed physiological changes similar to those in caloric restricted rodents, with the exception for IGF-1 and GH serum levels [109-111]. This suggests that the mechanisms linking early-life ER to cancer risk in animal experimental models of cancer cannot directly be extrapolated to humans.

## Future directions for human observational research

It seems that a negative energy balance in childhood and adolescence may impact on cancer occurring much later in life. However, the heterogeneity of observational studies to date makes it difficult to draw conclusions. This raises the question on how to proceed in this field. Molecular epidemiological approaches within existing studies may contribute to better insight into the mechanisms that may be at play. However, epidemiologic data regarding the mechanisms underlying an association between early-life ER and human site-specific cancer risk are scarce, because exposure to ER is rarely available in observational studies and few studies are large enough to allow for small subgroup analyses. In addition, tissues and molecular markers to investigate mechanisms are not commonly available. Tumor material, stored in pathology labs, can offer new opportunities for ongoing large-scale epidemiological studies since the tumors may provide molecular signatures of a carcinogenic process that started years ago.

Epigenetic changes are thought to be an early step in the carcinogenic process, typically environmental influences on epigenetics are most prominent during childhood and adolescence, the time frame of susceptibility to epigenetic/ transcriptional modulations that undergo establishment and maturation [157,158]. These epigenetic patterns can persist throughout life when occurring in stem cells [159]. Epigenetic markers can therefore be employed as a molecular signature to study how environmental exposures early in life may induce persistent epigenetic changes that influence methylation patterns in cancer occurring much later in life [160]. Hypermethylation through the CpG island methylator phenotype (CIMP) in the promotor region of specific cancer-related genes is considered an early event in carcinogenesis[161,162] and associations between early-life indicators of energy balance and CIMP in CRC may exist in particular. ER in adolescence has been inversely associated with CRC CIMP phenotype [163] which suggests that exposure to a transient environmental condition during this period of life can lead to sustained epigenetic modifications that impact cancer risk in adult life. Early-life ER has also been inversely associated with the risk of having a colorectal tumor characterized by IGFBP methylation [64]. Even though these types of molecular epidemiologic data are scarce, they are supportive of an inverse association between early-life ER and the risk of colorectal cancer. Therefore, replication of these studies and extension to other sites and mechanisms are needed to further substantiate the evidence.

## Conclusion

In general, it seems that severe transient ER in the absence of a nutritious diet is associated with increased cancer risk in the breast (for ER exposure at adolescent age) and prostate. Evidence for associations between severe transient ER early in life and risk at other cancer sites is limited. In the meta-analysis of the prospective cohort studies, the duration, rather than severity of exposure to early-life ER, seems to positively influence relative risk estimates. Results should be interpreted with caution due to the limited number of studies and difficulty in disentangling duration, severity and geographical setting of the exposure. For exposure to less severe ER, a decreased association with cancer risk is generally observed, although this is derived only from ecological studies. This raises the question on how to proceed in this field. Molecular epidemiological approaches within existing studies may contribute to explain in part the variation in disease risk across sites providing better insight into the mechanisms that might be at play.



## Acknowledgments

This work was supported by Wereld Kanker Onderzoek Fonds Nederland (WCRF-NL), as part of the World Cancer Research Fund International grant program (grant numbers 2012/618 and 2013/973 to MPW).

We thank Dr. H. Hoofs for assisting with the statistical software environment (version 3.1.2) R.

# References


1. Sell C. Caloric restriction and insulin-like growth factors in aging and cancer. Horm Metab Res. 2003;35(11-12):705-11. doi: 10.1055/s-2004-814156. PubMed PMID: 14710349.
2. Omodei D, Fontana L. Calorie restriction and prevention of age-associated chronic disease. FEBS Lett. 2011;585(11):1537-42. doi: 10.1016/j.febslet.2011.03.015. PubMed PMID: 21402069; PubMed Central PMCID: PMC3439843.
3. Speakman JR, Mitchell SE. Caloric restriction. Mol Aspects Med. 2011;32(3):159-221. doi: 10.1016/j.mam.2011.07.001. PubMed PMID: 21840335.
4. Roberts SB, Speakman J. Update on human calorie restriction research. Adv Nutr. 2013;4(5):563-4. doi: 10.3945/an.113.004317. PubMed PMID: 24038258; PubMed Central PMCID: PMC3771150.
5. Ross MH, Bras G. Lasting influence of early caloric restriction on prevalence of neoplasms in the rat. J Natl Cancer Inst. 1971;47(5):1095-113. PubMed PMID: 4330798.
6. Engelman RW, Day NK, Good RA. Calorie intake during mammary development influences cancer risk: lasting inhibition of C3H/HeOu mammary tumorigenesis by peripubertal calorie restriction. Cancer Res. 1994;54(21):5724-30. PubMed PMID: 7923222.
7. Koupil I, Plavinskaja S, Parfenova N, Shestov DB, Danziger PD, Vagero D. Cancer mortality in women and men who survived the siege of Leningrad (1941-1944). International Journal of Cancer. 2009;124(6):1416-21. doi: Doi 10.1002/Ijc.24093. PubMed PMID: WOS:000263539600024.
8. Cleary MP, Grossmann ME. The manner in which calories are restricted impacts mammary tumor cancer prevention. J Carcinog. 2011;10:21. doi: 10.4103/1477-3163.85181. PubMed PMID: 22013391; PubMed Central PMCID: PMC3190408.
9. Nolen GA. Effect of various restricted dietary regimens on the growth, health and longevity of albino rats. J Nutr. 1972;102(11):1477-93. PubMed PMID: 5081683.
10. Sylvester PW, Aylsworth CF, Van Vugt DA, Meites J. Influence of underfeeding during the "critical period" or thereafter on carcinogen-induced mammary tumors in rats. Cancer Res. 1982;42(12):4943-7. PubMed PMID: 6814746.
11. Cheney KE, Liu RK, Smith GS, Meredith PJ, Mickey MR, Walford RL. The effect of dietary restriction of varying duration on survival, tumor patterns, immune function, and body temperature in B10C3F1 female mice. J Gerontol. 1983;38(4):420-30. PubMed PMID: 6306089.
12. Kritchevsky D, Klurfeld DM. Caloric effects in experimental mammary tumorigenesis. Am J Clin Nutr. 1987;45(1 Suppl):236-42. PubMed PMID: 3799515.
13. Klurfeld DM, Welch CB, Lloyd LM, Kritchevsky D. Inhibition of DMBA-induced mammary tumorigenesis by caloric restriction in rats fed high-fat diets. Int J Cancer. 1989;43(5):922-5. PubMed PMID: 2497075.
14. Klurfeld DM, Welch CB, Davis MJ, Kritchevsky D. Determination of degree of energy restriction necessary to reduce DMBA-induced mammary tumorigenesis in rats during the promotion phase. J Nutr. 1989;119(2):286-91. PubMed PMID: 2493082.
15. Kritchevsky D, Welch CB, Klurfeld DM. Response of mammary tumors to caloric restriction for different time periods during the promotion phase. Nutr Cancer. 1989;12(3):259-69. doi: 10.1080/01635588909514025. PubMed PMID: 2505241.

16. Kumar SP, Roy SJ, Tokumo K, Reddy BS. Effect of different levels of calorie restriction on azoxymethane-induced colon carcinogenesis in male F344 rats. Cancer Res. 1990;50(18):5761-6. PubMed PMID: 2393850.
17. Roebuck BD, Baumgartner KJ, MacMillan DL. Caloric restriction and intervention in pancreatic carcinogenesis in the rat. Cancer Res. 1993;53(1):46-52. PubMed PMID: 8416749.
18. Dirx MJ, Zeegers MP, Dagnelie PC, van den Bogaard T, van den Brandt PA. Energy restriction and the risk of spontaneous mammary tumors in mice: a meta-analysis. Int J Cancer. 2003;106(5):766-70. doi: 10.1002/ijc.11277. PubMed PMID: 12866038.
19. Dogan S, Rogozina OP, Lokshin AE, Grande JP, Cleary MP. Effects of chronic vs. intermittent calorie restriction on mammary tumor incidence and serum adiponectin and leptin levels in MMTV-TGF-alpha mice at different ages. Oncol Lett. 2010;1(1):167-76. doi: 10.3892/ol_00000031. PubMed PMID: 22966277; PubMed Central PMCID: PMC3436387.
20. Klebanov S. Can short-term dietary restriction and fasting have a long-term anticarcinogenic effect? Interdiscip Top Gerontol. 2007;35:176-92. doi: 10.1159/000096562. PubMed PMID: 17063039.
21. Cleary MP, Hu X, Grossmann ME, Juneja SC, Dogan S, Grande JP, et al. Prevention of mammary tumorigenesis by intermittent caloric restriction: does caloric intake during refeeding modulate the response? Exp Biol Med (Maywood). 2007;232(1):70-80. PubMed PMID: 17202587.
22. Elias SG, Peeters PH, Grobbee DE, van Noord PA. Transient caloric restriction and cancer risk (The Netherlands). Cancer Causes Control. 2007;18(1):1-5. doi: 10.1007/s10552-006-0080-0. PubMed PMID: 17186418; PubMed Central PMCID: PMC1764866.
23. Mizuno NK, Rogozina OP, Seppanen CM, Liao DJ, Cleary MP, Grossmann ME. Combination of intermittent calorie restriction and eicosapentaenoic acid for inhibition of mammary tumors. Cancer Prev Res (Phila). 2013;6(6):540-7. doi: 10.1158/1940-6207.CAPR-13-0033. PubMed PMID: 23550153; PubMed Central PMCID: PMC4296517.
24. Lv M, Zhu X, Wang H, Wang F, Guan W. Roles of caloric restriction, ketogenic diet and intermittent fasting during initiation, progression and metastasis of cancer in animal models: a systematic review and meta-analysis. Plos One. 2014;9(12):e115147. doi: 10.1371/journal.pone.0115147. PubMed PMID: 25502434; PubMed Central PMCID: PMC4263749.
25. Fair AM, Montgomery K. Energy balance, physical activity, and cancer risk. Methods Mol Biol. 2009;472:57-88. doi: 10.1007/978-1-60327-492-0_3. PubMed PMID: 19107429.
26. Walford RL, Harris SB, Gunion MW. The calorically restricted low-fat nutrient-dense diet in Biosphere 2 significantly lowers blood glucose, total leukocyte count, cholesterol, and blood pressure in humans. Proc Natl Acad Sci U S A. 1992;89(23):11533-7. PubMed PMID: 1454844; PubMed Central PMCID: PMC50586.
27. Walford RL, Mock D, MacCallum T, Laseter JL. Physiologic changes in humans subjected to severe, selective calorie restriction for two years in biosphere 2: health, aging, and toxicological perspectives. Toxicol Sci. 1999;52(2 Suppl):61-5. PubMed PMID: 10630592.
28. Walford RL, Mock D, Verdery R, MacCallum T. Calorie restriction in biosphere 2: alterations in physiologic, hematologic, hormonal, and biochemical parameters in humans restricted for a 2-year period. J Gerontol A Biol Sci Med Sci. 2002;57(6):B211-24. PubMed PMID: 12023257.
29. Vin-Raviv N, Dekel R, Barchana M, Linn S, Keinan-Boker L. World War II-related post-traumatic stress disorder and breast cancer risk among Israeli women: a case-control study. Int Psychogeriatr. 2014;26(3):499-508. doi: 10.1017/S1041610213002081. PubMed PMID: 24290080.

30. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med. 2009;3(3):e123-30. PubMed PMID: 21603045; PubMed Central PMCID: PMCPMC3090117.
31. Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. 2009. Epub Available from: URL: http://www.ohri.ca/programs/clinical_epidemiology/ oxford.htm [cited 2009 Oct 19].
32. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36(3):1-48. PubMed PMID: WOS:000281593200001.
33. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58. doi: 10.1002/sim.1186. PubMed PMID: 12111919.
34. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PubMed PMID: 12958120; PubMed Central PMCID: PMC192859.
35. Bown MJ, Sutton AJ. Quality control in systematic reviews and meta-analyses. Eur J Vasc Endovasc Surg. 2010;40(5):669-77. doi: 10.1016/j.ejvs.2010.07.011. PubMed PMID: 20732826.
36. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. British Medical Journal. 2011;343. doi: ARTN d400210.1136/bmj.d4002. PubMed PMID: WOS:000293169000008.
37. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34. PubMed PMID: 9310563; PubMed Central PMCID: PMC2127453.
38. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Statistics in Medicine. 2002;21(11):1559-73. doi: 10.1002/sim.1187. PubMed PMID: WOS:000176016900006.
39. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey LH. Energy restriction early in life and colon carcinoma risk: results of The Netherlands Cohort Study after 7.3 years of follow-up. Cancer. 2003;97(1):46-55. doi: 10.1002/cncr.11052. PubMed PMID: 12491504.
40. Hughes LA, van den Brandt PA, Goldbohm RA, de Goeij AF, de Bruine AP, van Engeland M, et al. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. Int J Epidemiol. 2010;39(5):1333-44. doi: 10.1093/ije/dyq062. PubMed PMID: 20427463.
41. Elias SG, Peeters PHM, Grobbee DE, Noord PAHv. Breast Cancer Risk After Caloric Restriction During the 1944-1945 Dutch Famine. JNCI Journal of the National Cancer Institute. 2004;96(7):539-46. doi: 10.1093/jnci/djh087.
42. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey LH. Diet in adolescence and the risk of breast cancer: results of the Netherlands Cohort Study. Cancer Causes Control. 1999;10(3):189-99. PubMed PMID: 10454064.
43. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey LH. Energy restriction in childhood and adolescence and risk of prostate cancer: results from the Netherlands Cohort Study. Am J Epidemiol. 2001;154(6):530-7. PubMed PMID: 11549558.

44. Robsahm TE, Tretli S. Breast cancer incidence in food- vs non-food-producing areas in Norway: possible beneficial effects of World War II. Br J Cancer. 2002;86(3):362-6. doi: 10.1038/sj.bjc.6600084. PubMed PMID: 11875700; PubMed Central PMCID: PMC2375214.
45. Fentiman IS, Allen DS, Ellison GT. The impact of the Occupation of Guernsey 1940-1945 on breast cancer risk factors and incidence. Int J Clin Pract. 2007;61(6):937-43. doi: 10.1111/j.1742-1241.2007.01288.x. PubMed PMID: 17504356.
46. Heinen MM, Verhage BA, Goldbohm RA, Lumey LH, van den Brandt PA. Physical activity, energy restriction, and the risk of pancreatic cancer: a prospective study in the Netherlands. Am J Clin Nutr. 2011;94(5):1314-23. doi: 10.3945/ajcn.110.007542. PubMed PMID: 21955648.
47. Da Li Q, Li H, Li FJ, Wang MS, Li ZJ, Han J, et al. Nutrition deficiency increases the risk of stomach cancer mortality. Bmc Cancer. 2012;12. doi: Artn 315 Doi 10.1186/1471-2407-12-315. PubMed PMID: WOS:000308808000001.
48. Keinan-Boker L, Vin-Raviv N, Liphshitz I, Linn S, Barchana M. Cancer incidence in Israeli Jewish survivors of World War II. J Natl Cancer Inst. 2009;101(21):1489-500. doi: 10.1093/jnci/djp327. PubMed PMID: 19861305.
49. Vin-Raviv N, Barchana M, Linn S, Keinan-Boker L. Severe caloric restriction in young women during World War II and subsequent breast cancer risk. Int J Clin Pract. 2012;66(10):948-58. doi: 10.1111/j.1742-1241.2012.02966.x. PubMed PMID: 22994329.
50. Moller H. Decreased testicular cancer risk in men born in wartime. J Natl Cancer Inst. 1989;81(21):1668-9. PubMed PMID: 2795697.
51. Wanderas EH, Tretli S, Fossa SD. Trends in incidence of testicular cancer in Norway 1955-1992. Eur J Cancer. 1995;31A(12):2044-8. PubMed PMID: 8562163.
52. Bergstrom R, Adami HO, Mohner M, Zatonski W, Storm H, Ekbom A, et al. Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. J Natl Cancer Inst. 1996;88(11):727-33. PubMed PMID: 8637026.
53. Tretli S, Gaard M. Lifestyle changes during adolescence and risk of breast cancer: an ecologic study of the effect of World War II in Norway. Cancer Causes Control. 1996;7(5):507-12. PubMed PMID: 8877047.
54. Thorn M, Bergstrom R, Kressner U, Sparen P, Zack M, Ekbom A. Trends in colorectal cancer incidence in Sweden 1959-93 by gender, localization, time period, and birth cohort. Cancer Causes Control. 1998;9(2):145-52. PubMed PMID: 9578291.
55. Becker N, Muscat JE, Wynder EL. Cancer mortality in the United States and Germany. J Cancer Res Clin Oncol. 2001;127(5):293-300. PubMed PMID: 11355144.
56. Moller H. Trends in incidence of testicular cancer and prostate cancer in Denmark. Hum Reprod. 2001;16(5):1007-11. PubMed PMID: 11331652.
57. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. Eur J Cancer Prev. 2002;11(5):489-95. PubMed PMID: 12394247.
58. Richiardi L BR, Adami HO, Torrång A, Barlow L, Hakulinen T, Rahu M, Stengrevics A, Storm H, Tretli S, Kurtinaitis J, Tyczynski JE, Akre O. Testicular cancer incidence in eight northern European countries: secular and recent trends. Cancer Epidemiol Biomarkers Prev. 2004;(13(12)):2157-66.

59. Svensson E, Moller B, Tretli S, Barlow L, Engholm G, Pukkala E, et al. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. Cancer Causes Control. 2005;16(3):215-23. doi: 10.1007/s10552-004-3073-x. PubMed PMID: 15947873.
60. Jacobsen R, Moller H, Thoresen SO, Pukkala E, Kjaer SK, Johansen C. Trends in testicular cancer incidence in the Nordic countries, focusing on the recent decrease in Denmark. Int J Androl. 2006;29(1):199-204. doi: 10.1111/j.1365-2605.2005.00605.x. PubMed PMID: 16371112.
61. Borsoi L, Kunze U, Kunze M, Groman E, Kundi M. Trends in mortality and mean age at death from lung cancer in Austria (1975-2007). Cancer Epidemiol. 2011;35(2):120-5. doi: 10.1016/j.canep.2010.06.018. PubMed PMID: 20673657.
62. Chauvenet M, Cottet V, Lepage C, Jooste V, Faivre J, Bouvier AM. Trends in colorectal cancer incidence: a period and birth-cohort analysis in a well-defined French population. BMC Cancer. 2011;11:282. doi: 10.1186/1471-2407-11-282. PubMed PMID: 21718477; PubMed Central PMCID: PMC3149029.
63. Hughes LA, Simons CC, van den Brandt PA, Goldbohm RA, van Engeland M, Weijenberg MP. Body size and colorectal cancer risk after 16.3 years of follow-up: an analysis from the Netherlands Cohort Study. Am J Epidemiol. 2011;174(10):1127-39. doi: 10.1093/aje/kwr247. PubMed PMID: 21984660.
64. Schouten LJ, van Dijk BA, Lumey LH, Goldbohm RA, van den Brandt PA. Energy restriction during childhood and early adulthood and ovarian cancer risk. Plos One. 2011;6(11):e27960. doi: 10.1371/journal.pone.0027960. PubMed PMID: 22132180; PubMed Central PMCID: PMC3223198.
65. Angell-Andersen E, Tretli S, Bjerknes R, Forsen T, Sorensen TI, Eriksson JG, et al. The association between nutritional conditions during World War II and childhood anthropometric variables in the Nordic countries. Ann Hum Biol. 2004;31(3):342-55. doi: 10.1080/03014460410001685304. PubMed PMID: 15204349.
66. Mahabir S. Association between diet during preadolescence and adolescence and risk for breast cancer during adulthood. J Adolesc Health. 2013;52(5 Suppl):S30-5. doi: 10.1016/j.jadohealth.2012.08.008. PubMed PMID: 23298994; PubMed Central PMCID: PMC3622736.
67. van Noord PA. Breast cancer and the brain: a neurodevelopmental hypothesis to explain the opposing effects of caloric deprivation during the Dutch famine of 1944-1945 on breast cancer and its risk factors. J Nutr. 2004;134(12 Suppl):3399S-406S. PubMed PMID: 15570045.
68. Land CE, Tokunaga M, Koyama K, Soda M, Preston DL, Nishimori I, et al. Incidence of female breast cancer among atomic bomb survivors, Hiroshima and Nagasaki, 1950-1990. Radiat Res. 2003;160(6):707-17. doi: Doi 10.1667/Rr3082. PubMed PMID: WOS:000187079200012.
69. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. Breast Cancer Res. 2005;7(1):21-32. doi: 10.1186/bcr970. PubMed PMID: 15642178; PubMed Central PMCID: PMCPMC1064116.
70. Papadopoulos FC, Pantziaras I, Lagiou P, Brandt L, Ekselius L, Ekbom A. Age at onset of anorexia nervosa and breast cancer risk. Eur J Cancer Prev. 2009;18(3):207-11. doi: 10.1097/CEJ.0b013e32831bc558. PubMed PMID: 19491607.
71. Hilakivi-Clarke L. Nutritional modulation of terminal end buds: its relevance to breast cancer prevention. Curr Cancer Drug Targets. 2007;7(5):465-74. PubMed PMID: 17691906.
72. Brundtland GH, Liestol K, Walloe L. Height, weight and menarcheal age of Oslo schoolchildren during the last 60 years. Ann Hum Biol. 1980;7(4):307-22. PubMed PMID: 7436344.

73. Frankel S, Gunnell DJ, Peters TJ, Maynard M, Davey Smith G. Childhood energy intake and adult mortality from cancer: the Boyd Orr Cohort Study. BMJ. 1998;316(7130):499-504. PubMed PMID: 9501710; PubMed Central PMCID: PMC2665640.
74. Dols M vAA. Voedselvoorziening in Nederland tijdens en onmiddellijk na den Tweeden Wereldoorlog 1940-1945 (Food supply during and after the second World War 1940-1945 in the Netherlands [in Dutch]). Voeding. 1946;(6):193-207.
75. Burger G SH, Drummond J. Malnutrition and starvation in Western Netherlands, September 1944-45. Part I and II. The Hague: General State Printing Office. 1948.
76. Hansen OG. Food conditions in Norway during the war, 1939-45. Proc Nutr Soc. 1947;5:263-70.
77. Strøm A. Examination Into the Diet of Norwegian Families During the War-years 1942-1945: Exp.; 1948.
78. Bennett MK. Food and agriculture in the Soviet Union, 1917-48. J Polit Econ. 1949;57:185-98.
79. Elias SG, Keinan-Boker L, Peeters PH, Van Gils CH, Kaaks R, Grobbee DE, et al. Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. Int J Cancer. 2004;108(4):628-30. doi: 10.1002/ijc.11584. PubMed PMID: 14696131.
80. Hursting SD, Smith SM, Lashinger LM, Harvey AE, Perkins SN. Calories and carcinogenesis: lessons learned from 30 years of calorie restriction research. Carcinogenesis. 2010;31(1):83-9. doi: 10.1093/carcin/bgp280. PubMed PMID: 19969554.
81. Anisimov VN, Bartke A. The key role of growth hormone-insulin-IGF-1 signaling in aging and cancer. Crit Rev Oncol Hematol. 2013;87(3):201-23. doi: 10.1016/j.critrevonc.2013.01.005. PubMed PMID: 23434537.
82. McCarty CA. Up-regulation of IGF binding protein-1 as an anticarcinogenic strategy: relevance to caloric restriction, exercise and insulin sensitivty. Medical Hypothese. 1997.
83. Martin B, Golden E, Carlson OD, Egan JM, Mattson MP, Maudsley S. Caloric restriction: Impact upon pituitary function and reproduction. Ageing Res Rev. 2008;7(3):209-24. doi: 10.1016/j.arr.2008.01.002. PubMed PMID: WOS:000260621200005.
84. Okasha M, Gunnell D, Holly J, Davey Smith G. Childhood growth and adult cancer. Best Pract Res Clin Endocrinol Metab. 2002;16(2):225-41. doi: 10.1053/beem.2002.0204. PubMed PMID: 12064890.
85. Thissen JP, Ketelslegers JM, Underwood LE. Nutritional regulation of the insulin-like growth factors. Endocr Rev. 1994;15(1):80-101. doi: 10.1210/edrv-15-1-80. PubMed PMID: 8156941.
86. Soeters MR, Soeters PB, Schooneman MG, Houten SM, Romijn JA. Adaptive reciprocity of lipid and glucose metabolism in human short-term starvation. Am J Physiol Endocrinol Metab. 2012;303(12):E1397-407. doi: 10.1152/ajpendo.00397.2012. PubMed PMID: 23074240.
87. Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Jr S, et al. Epigenetic programming by maternal behavior. Nat Neurosci. 2004;7(8):847-54. doi: 10.1038/nn1276. PubMed PMID: WOS:000222930800015.
88. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestart ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. P Natl Acad Sci USA. 2005;102(30):10604-9. doi: 10.1073/pnas.0500398102. PubMed PMID: WOS:000230853300037.
89. De Assis S, Hilakivi-Clarke L. Timing of dietary estrogenic exposures and breast cancer risk. Ann N Y Acad Sci. 2006;1089:14-35. doi: 10.1196/annals.1386.039. PubMed PMID: 17261753.

# Supplementary data

### S1 Table. Prisma 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| TITLE | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| ABSTRACT | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 1-2 |
| INTRODUCTION | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 3-4 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 4 |
| METHODS | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | 4 |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 5 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 5-6 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 5 |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 5-8 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 5-6 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 5-6 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 6 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 5 |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2 for each meta-analysis.) | 6-8 |
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 8 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 8 |

The table Continues on the next page

| | | | |
|---|---|---|---|
| RESULTS | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 8-9 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 8-13 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 13-14 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 14-19 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 14-19 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 13-14 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | 19-23 |
| DISCUSSION | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 24-29 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 29-30 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 31 |
| FUNDING | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 32 |

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

**S2 Table: Overview of characteristics of ecological studies describing birth cohort trends in cancer incidence and cancer mortality possibly connected to energy restriction in the period encompassing World War II.**

| Study[a] | Country | Estimated ER[a] | N cases | (Poisson) model | Birth Cohorts |
|---|---|---|---|---|---|
| **Breast cancer incidence** | | | | | |
| *Tretli et al., 1996* | Norway | 20%[b] | 20,111 | Age-period-cohort model and exposure/sensitivity model | 1903-1953 |
| **Prostate cancer incidence** | | | | | |
| *Moller et al., 2001* | Denmark | 7%[b] | NA | Age-period-cohort model | 1858-1948 |
| **Colorectal cancer incidence** | | | | | |
| *Thörn et al., 1998* | Sweden | 4%[b] | Colon cancer: 85,547; rectal cancer: 49,096 | Age-period-cohort model by gender | 1875-1974 |
| *Svensson et al., 2002* | Norway | 20%[b] | Men: 32,981; women: 32,812 | Age-period-cohort model by gender | 1874-1953 |
| *Svensson et al., 2005* | Norway (NO), Sweden (SE), Denmark (DK), Finland (FI), Iceland (IC), Estonia (ES) | NO: 20%; FI: 17%; DK: 7%; SE: 4%; IC: NA; ES; NA[b] | NO: 61,836 ; FI: 35,003; DK: 88,832 ; SE:119,416; IC: 1,762; ES: 10,224 | Age-period-cohort model by gender | 1874-1953 |
| **Testicular cancer incidence** | | | | | |
| *Wanderås et al., 1995* | Norway | 20%[b] | 3,927 | Age-period-cohort model | 1916-1970 |
| *Bergström et al., 1996* | Denmark (DK), Norway (NO), Sweden (SE), East Germany (E-GER), Finland (FI), Poland (PL) | DK: 7%; NO: 20%; SE: 4%; E-GER: NA; FI: 17%; PL: NA[b] | DK: 6,352; NO: 3,111; SE: 3,770; E-GER: 10,051; FI: 1,174; PL: 6,450 | Age-period-cohort model | 1860-1969 (DK, fewer birth cohorts in other countries) |
| *Moller et al. 1989; Moller et al., 2001* | Denmark | 7%[b] | NA | Age-period-cohort model | 1888-1978 |
| *Richiardi et al., 2004* | Denmark (DK), Estonia (ES), Finland (FI), Latvia (LAT), Lithuania (LIT), Norway (NO), Poland (PL), Sweden (SE) | DK: 7%; ES: NA; FI: 17%; LAT: NA; LIT: NA; NO: 20%; PL: NA; SE: 4%[b] | DK: 9,138; ES: 293; FI: 1,842; LAT: 392; LIT: 259; NO: 4,888; PL: 4,388; SE: 5,830 | Age-period-cohort model for DK, FI, NO and SE, which had a sufficiently long period of registration and a large number of cases | 1885-1980 (DK, fewer birth cohorts in other countries) |

The table Continues on the next page

| | | | | | |
|---|---|---|---|---|---|
| *Jacobsen et al., 2006* | Denmark (DK), Finland (FI), Norway (NO), Sweden (SE) | DK: 7%; FI: 17%; NO: 20%; SE: 4%[b] | NA | Age-standardized rates and incidence rates over age by birth cohort | 1938-1983 |
| **Lung cancer mortality** | | | | | |
| *Borsoi et al., 2011* | Austria | NA | NA | Age-period-cohort model by gender | 1880-1960+ |
| **Site-specific cancer mortality** | | | | | |
| *Becker al., 2001* | West-Germany versus United States of America (USA) | ~40-50%[c] | NA | Age-standardized mortality rates by gender | NA |

Abbreviations: CRC, colorectal cancer; NA, not available.

[a] The main results from these studies are described in the manuscript and full references are included at the end of the manuscript.

[b] Angell-Andersen et al. (2004). Ann Hum Biology, 31(3): 342-355. DOI:10.1080/03014460410001685304. In addition, Lund Nilsen and Vatten (2001) report intakes of 2700 kcal/day during World War 2 versus 3475 kcal/day pre-war in Norwegian families (Br J Cancer, 85(7): 959-961. DOI: 10.1054/ bjoc.2001.1946).

[c] Becker et al., 2001 give the following description: while in 1937 the average calorie demand per capita was assessed to be 2,750 kcal per day, the food allocated by the food ration cards led to a decreasing calorie supply, reaching a nadir in 1945 with estimated 1,412-1,600 kcal per capita per day.

**S3 Table: Study characteristics concerning the duration and caloric intake as reported by the included cohort studies.**

| Cohort[a] | Country | Historical event | Intervention period in months (daily caloric intake) | Additional notes[a] |
|---|---|---|---|---|
| Elias et al, 2004 | The Netherlands | Dutch Hunger Winter | 6 (<700 from January 1945 onwards) | The diet remained nutritionally balanced, but the population living in western urban areas of the Netherlands experienced a rationing of <700 kilocalories per capita per day from January 1945 onwards *(Burger et al., 1948; Dols et al., 1946)*. |
| Fentiman et al, 2007 | England | Occupation of Guernsey | 11 (1200-1660) | From June 1944 onwards, the island's food supplies were severely limited to approximately 1200 kilocalories per capita per day, which lasted until May 1945, although some Red Cross parcels brought relieve (~460 kilocalories) from December 1944 onwards *(Fentiman et al., 2007)*. |
| Keinan-boker et al, 2009 | Israël | Holocaust | 72 (220-800) | Jews experienced a long-term and severe energy restriction due to being interred in concentration camps and ghettos at the beginning of World War 2, often resulting in malnutrition and associated clinical manifestations, *e.g.* rickets, night blindness, anemia, and scurvy *(Keinan-Boker et al., 2009)*. |
| Koupil et al, 2009 | Russia | Siege of Leningrad | 28 (300 in winter 1941-42) | The siege of Leningrad lasted from September 1941 to January 1944 (28 months), leading to an acute food shortage and severe malnutrition, with its peak in the winter of 1941-42 when ~300 kilocalories per capita per day were provided *(Koupil et al., 2009)*. |
| Robsahm et al, 2009 | Norway | World War II | 60 (20% restricted) | The diet remained nutritionally balanced during the war years in Norway, but was 20% restricted (*Angell-w* |

The table Continues on the next page

| | | | | |
|---|---|---|---|---|
| Li et al, 2012 | China (Zhaoyuan County) | Chinese famine during "Great Leap Forward" | >36 with a 24-month peak (-) | The famine affected all of China, but its severity and duration, including the beginning and end, varied across regions, and, therefore, the famine is difficult to define for any given area with precision; the Chinese famine followed the already widespread presence of chronic undernutrition *(Huang et al. 2010; J Nutr, 140:1874–1878. DOI:10.3945/jn.110.121293)*. The severity of exposure for a particular county was estimated by the cohort size shrinkage estimate, which took into account the mean cohort size of a person born during the famine and of a person born immediately before and after the famine. This shrinkage estimate indicated that Zhaoyuan County had a value near the highest of famine indexes for 35 counties; in addition, the duration of the famine in Zhaoyuan County was estimated to be longer than average *(Huang et al. 2010; J Nutr, 140:1874–1878. DOI:10.3945/jn.110.121293; Li et al., 2012)*. |
| Dirx et al, 1999, Dirx et al, 2001, Hughes et al, 2010, Heinen et al, 2011, Schouten et al, 2011 | The Netherlands | Dutch Hunger Winter | 5 (<700 from January 1945 onwards) | See *Elias et al., 2004.* The height of the Dutch famine has been described to have been from December 1944 until April 1945, lasting approximately 5 months (depending on the exact definition). |

[a] References not given in full are available in full in the manuscript.

**S4 Table: Qualitative assessment of included cohort studies according to the quality subscales of the Newcatle-Ottowa scale.**

| Cohort | Quality assessment score: Total (max. 8 points) | Score for subscale: Selection (max. 4 points) | Score for subscale: Comparability (max. 1 point) | Score for subscale: Outcome (max. 3 points) |
|---|---|---|---|---|
| Dirx et al, 1999 | 7 | 3 | 1 | 3 |
| Dirx et al, 2001 | 7 | 3 | 1 | 3 |
| Elias et al, 2004 | 6 | 2 | 1 | 3 |
| Fentiman et al, 2007 | 6 | 3 | 1 | 2 |
| Keinan-Boker et al, 2009 | 7 | 3 | 1 | 3 |
| Koupil et al, 2009 | 6 | 3 | 1 | 2 |
| Robsahm et al, 2009 | 6 | 3 | 1 | 2 |
| Hughes et al, 2010 | 7 | 3 | 1 | 3 |
| Heinen et al, 2011 | 7 | 3 | 1 | 3 |
| Schouten et al, 2011 | 7 | 3 | 1 | 3 |
| Li et al, 2012 | 6 | 3 | 1 | 2 |

Abbreviations: max., maximum.

**S5 Table: Overview of cohorts investigating transient energy restriction in early-life and the relative risk or mortality of breast cancer in females.**

| Study cohort | Reference | Measured exposure contrast | Outcome | Estimated duration | Estimated energy | PY | | N cases | | Birth Cohorta | Age at exposure in yearsa | Relative risk | 95% CI | | Model |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (Exposed vs. reference) | | of exposure in months | restriction in kcal/day | Exposed | Reference | Exposed | Reference | | | | lower bound | upper bound | |
| NLCS | Dirx et al., 1999 | Lived in Western city vs. non-Western area | Breast cancer incidence | 5b | <700 (after January 1945) | 2,420 | 4,809 | 239 | 418 | Birth Cohort 1916-1932 | 11-28 | 1.10c | 0.90 | 1.40 | Multivariabled |
| NLCS | Dirx et al., 1999 | Lived in Western city vs. non-Western area | Breast cancer incidence | 5b | <700 (after January 1945) | 100 | 100 | 2 | 8 | - | Before growth spurt | Not estimable (too few cases) | - | - | Multivariabled |
| NLCS | Dirx et al., 1999 | Lived in Western city vs. non-Western area | Breast cancer incidence | 5b | <700 (after January 1945) | 604 | 1,308 | 60 | 102 | - | During growth spurt (2 yr< menarche<1 yr) | 1.20 | 0.80 | 2.00 | Multivariabled |
| NLCS | Dirx et al., 1999 | Lived in Western city vs. non-Western area | Breast cancer incidence | 5b | <700 (after January 1945) | 1,738 | 3,348 | 182 | 332 | - | After growth spurt | 1.10 | 0.90 | 1.40 | Multivariabled |
| DOM Study | Ellias et al., 2004 | Hunger vs. no hunger (self-reports) | Breast cancer incidence | 6 | <700 (after January 1945) | 21,415 | 109,502 | 79 | 265 | Birth Cohort 1911-1941 | 2-33 | 1.48c | 1.09 | 2.01 | Multivariablee |
| DOM Study | Ellias et al., 2004 | Hunger vs. no hunger (self-reports) | Breast cancer incidence | 6 | <700 (after January 1945) | 2,465 | 32,818 | 14 | 77 | Birth Cohort 1935-1941 | 2-9 | 2.01 | 0.92 | 4.41 | Multivariablee |
| DOM Study | Ellias et al., 2004 | Hunger vs. no hunger (self-reports) | Breast cancer incidence | 6 | <700 (after January 1945) | 6,764 | 42,265 | 23 | 91 | Birth Cohort 1926-1934 | 10-18 | 1.55 | 0.89 | 2.73 | Multivariablee |
| DOM Study | Ellias et al., 2004 | Hunger vs. no hunger (self-reports) | Breast cancer incidence | 6 | <700 (after January 1945) | 12,186 | 34,419 | 42 | 97 | Birth Cohort ≤1925 | ≥19 | 1.18 | 0.77 | 1.80 | Multivariablee |
| Guernsey Study | Fentiman et al., 2007 | Stayed vs. evacuated | Breast cancer incidence | 11 | 1200-1660 | 20,438 | 15,294 | 37 | 60 | Birth cohort ≤1925-1946 | In utero to ≥19 | 0.91c | 0.60 | 1.37 | Multivariablef |
| Guernsey Study | Fentiman et al., 2007 | Stayed vs. evacuated | Breast cancer incidence | 11 | 1200-1660 | 8,336 | 8,589 | 14 | 23 | Birth Cohort 1935-1946 | In utero-9 | 0.68 | 0.35 | 1.34 | Multivariablef |
| Guernsey Study | Fentiman et al., 2007 | Stayed vs. evacuated | Breast cancer incidence | 11 | 1200-1660 | 2,847 | 8,267 | 10 | 22 | Birth Cohort 1926-1934 | 10-18 | 1.26 | 0.59 | 2.66 | Multivariablef |
| Guernsey Study | Fentiman et al., 2007 | Stayed vs. evacuated | Breast cancer incidence | 11 | 1200-1660 | 4,111 | 3,581 | 13 | 15 | Birth Cohort ≤1925 | ≥19 | 0.68 | 0.32 | 1.46 | Multivariablef |
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | 429,301 | 17,647 | 896 | 374 | Birth Cohort 1940-1945 | In utero-5 | 2.44 | 1.46 | 4.06 | Poisson (age and period) |
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | 377,733 | 40,244 | 1,124 | 683 | Birth Cohort 1935-1939 | In utero-10 | 1.63 | 1.29 | 2.06 | Poisson (age and period) |

The table Continues on the next page

S5 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | 387,166 | 112,007 | 1,328 | 829 | Birth Cohort 1930-1934 | 5-15 | 1.62 | 1.41 | 1.86 | Poisson (age and period) |
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | 492,743 | 132,929 | 1,761 | 1,538 | Birth Cohort 1925-1929 | 10-20 | 1.20 | 1.08 | 1.33 | Poisson (age and period) |
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | 519,343 | 161,125 | 1,996 | 1,773 | Birth Cohort 1920-1924 | 15-25 | 1.21 | 1.10 | 1.33 | Poisson (age and period) |
| Israeli Cohort | Keinan-boker et al., 2009 | Immigrated after the war vs. before the war | Breast cancer incidence | 72 | 220-800 | | | | | Pooled estimateg | | 1.46c | 1.20 | 1.77 | - |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Breast cancer mortality | 28 | 300 (winter 1941-42) | - | - | 12 | 6 | Birth Cohort 1916-1935 | 6-25 | 2.40c | 0.86 | 6.72 | Multivariableh |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Breast cancer mortality | 28 | 300 (winter 1941-42) | - | - | - | - | Birth Cohort 1939-1944 | 1-5 | 0.44 | 0.05 | 3.68 | Univariable (age) |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Breast cancer mortality | 28 | 300 (winter 1941-42) | - | - | - | - | Birth Cohort 1926-1931 | 10-18 | 9.70 | 1.10 | 84.5 | Univariable (age) |

Abbreviations: CI, confidence interval; NLCS, Netherlands Cohort Study; PY, person-years; US-USSR LP, US-USSR Lipid Program.

a In order to facilitate comparison between studies, birth cohort was derived from age at exposure if not mentioned in the paper and vice versa, and may therefore be an approximation.

b Same exposure as in Elias et al., 2004, though the height of the famine was considered

c Relative risk ratio or hazard ratio taken forward to the meta-analysis. If a cohort reported relative risk ratios or hazard ratios for multiple birth cohorts, a pooled estimate was calculated for these birth cohorts and the pooled estimate was taken forward in the meta-analysis.

d Hazard ratio adjusted for age, age at menopause, parity, age at first birth, maternal breast cancer, breast cancer in sister(s), benign breast disease, alcohol use, energy consumption, education, height and age at menarche.

e Hazard ratio adjusted for age at screening examination and age at screening examination squared, body mass index, height, socioeconomic status, age at menarche, parity, age at birth of first child and family history of breast cancer (first-degree relative).

f Hazard ratio adjusted for age at entry, age at menarche, age first birth/nulliparity, height and body mass index.

g Even though heterogeneity in estimates between birth cohorts was >50% for some cancer endpoints studied by Keinan-Boker et al., all birth cohorts were used in the pooling, because no specific birth cohort introduced heterogeneity across endpoints as based on the leave-one-out estimates of heterogeneity and because the pooled estimate did not materially differ when excluding the specific birth cohort introducing heterogeneity in analyses for a certain cancer endpoint.

h Hazard ratio adjusted for exact age at examination and social characteristics (education, occupation, ethnic group/nationality and marital status), behavioral characteristics (smoking, alcohol consumption) and adult stature (adult height and bodymass index).

Supplemental table 5 continious on the next page

**S6 Table: Overview of cohorts on the association between transient energy restriction in early-life and the relative risk or mortality of prostate cancer in males.**

| Study cohort | Reference | Measured exposure contrast | Outcome | Estimated duration of exposure in months | Estimated energy restriction in kcal/day | PY | | N cases | | Birth Cohorta | Age at exposure in yearsa | Relative risk | 95% CI | | Model |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (Exposed vs. reference) | | | | Exposed | Reference | Exposed | Reference | | | | lower bound | upper bound | |
| NLCS | Dirx et al., 2001 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 5 | <700 (from January 1945 onwards) | 2,090 | 4,914 | 168 | 363 | Birth Cohort 1916 -1932 | 11-28 | 1.15b | 0.80 | 1.31 | Multivariablec |
| NLCS | Dirx et al., 2001 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 5 | <700 (from January 1945 onwards) | - | - | - | - | Birth Cohort 1932 | Before growth spurt (11) | Not estimable (too few cases) | | | |
| NLCS | Dirx et al., 2001 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 5 | <700 (from January 1945 onwards) | 521 | 1,192 | 11 | 28 | Birth Cohort 1929-1932 | During growth spurt (12-15) | 1.01 | 0.46 | 2.22 | Multivariablec |
| NLCS | Dirx et al., 2001 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 5 | <700 (from January 1945 onwards) | 1,569 | 3,722 | 157 | 335 | Birth Cohort <1929 | After growth spurt (>15) | 1.03 | 0.79 | 1.35 | Multivariablec |
| Israeli Cohort | Keinan-Boker et al., 2009 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 72 | 220-800 | 338,512 | 39,129 | 308 | 278 | Birth Cohort 1935-1939 | In utero-10 | 1.08d | 0.76 | 1.53 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 72 | 220-800 | 326,174 | 102,040 | 529 | 397 | Birth Cohort 1930-1934 | 5-15 | 1.34d | 1.11 | 1.62 | Poisson (age and period) |

The table Continues on the next page

S6 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Israeli Cohort | Keinan-Boker et al., 2009 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 72 | 220-800 | 342,575 | 127,525 | 757 | 642 | Birth Cohort 1925-1929 | 10-20 | 1.20d | 1.04 | 1.37 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 72 | 220-800 | 397,834 | 158,052 | 1050 | 1,076 | Birth Cohort 1920-1924 | 15-25 | 1.02d | 0.91 | 1.13 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Hunger vs. no hunger (self-reports) | Prostate cancer incidence | 72 | 220-800 | | | | | Multivariablee | | 1.15b | 1.01 | 1.31 | |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Prostate cancer mortality | 28 | 300 (winter 1941-42) | - | - | 14 | 17 | Birth Cohort 1916-1935 | 6-25 | 1.43b | 0.70 | 2.92 | Multivariablef |

Abbreviations: CI, confidence interval; NLCS, Netherlands Cohort Study; PY, person-years; US-USSR LP, US-USSR Lipid Program.

a In order to facilitate comparison between studies, birth cohort was derived from age at exposure if not mentioned in the paper and vice versa, and may therefore be an approximation.

b Relative risk ratio or hazard ratio taken forward to the meta-analysis. If a cohort reported relative risk ratios or hazard ratios for multiple birth cohorts, a pooled estimate was calculated for these birth cohorts and the pooled estimate was taken forward to the meta-analysis.

c Model adjusted for age, prostate cancer in family, energy intake in 1986 (kcal/day), education, marital status, height, and β -cryptoxanthin intake in 1986 (μg/day).

d Model for prostate cancer cases diagnosed between 1991-2004 (after the introduction of prostate-specific antigen screening).

e Even though heterogeneity in estimates between birth cohorts was >50% for some cancer endpoints studied by Keinan-Boker et al., all birth cohorts were used in the pooling, because no specific birth cohort introduced heterogeneity across endpoints as based on the leave-one-out estimates of heterogeneity and because the pooled estimate did not materially differ when excluding the specific birth cohort introducing heterogeneity in analyses for a certain cancer endpoint.

f Model adjusted for exact age at examination and social characteristics (education, occupation, ethnic group/nationality and marital status), behavioral characteristics (smoking, alcohol consumption) and adult stature (adult height and body mass index).

Supplemental table 6 continuous on the next page

**S7 Table: Overview of cohorts on the association between transient energy restriction in early-life and the relative risk or mortality at sites other than prostate cancer risk in males and breast cancer risk in females.**

| Study cohort | Reference | Measured exposure contrast | Outcome | Estimated duration of exposure in months | Estimated energy restriction in kcal/day | PY | | N cases | | Birth Cohorta | Age at exposure | Relative risk | 95% CI | | Model |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (Exposed vs. reference) | | | | Exposed | Reference | Exposed | Reference | | in yearsa | | lower bound | upper bound | |
| **Males** | | | | | | | | | | | | | | | |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 377,733 | 40,244 | 388 | 213 | Birth Cohort 1935-1939 | In utero-10 | 1.75 | 1.19 | 2.59 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 387,166 | 112,007 | 658 | 359 | Birth Cohort 1930-1934 | 5-15 | 1.84 | 1.51 | 2.24 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 492,743 | 132,929 | 1,053 | 705 | Birth Cohort 1925-1929 | 10-20 | 1.56 | 1.37 | 1.78 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 519,343 | 161,125 | 1,682 | 1,341 | Birth Cohort 1920-1924 | 15-25 | 1.31 | 1.19 | 1.45 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | | | | | Pooled estimateb | - | 1.55c | 1.32 | 1.83 | |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Colorectal cancer mortality | 28 | 300 (winter 1941-42) | | | 17 | 34 | Birth Cohort 1916-1935 | 6-25 | 0.90c | 0.50 | 1.62 | Multivariabled |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 5,774 | 13,798 | 314 | 905 | Birth cohort 1916-1932 | 12-28 | 0.81c | 0.68 | 0.98 | Multivariablee |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | | | | | Birth Cohort 1932 | Before growth spurt (11) | Not estimable (no cases) | | | |

The table Continues on the next page

S7 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 1,042 | 2,383 | 21 | 87 | Birth Cohort 1929-1932 | During growth spurt (12-15) | 0.56 | 0.31 | 1.03 | Multivariablee |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 4,327 | 9,796 | 208 | 594 | Birth Cohort <1929 | After growth spurt (>15) | 0.76 | 0.61 | 0.94 | Multivariablee |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Stomach cancer mortality | 28 | 300 (winter 1941-42) | | | 37 | 60 | Birth Cohort 1916-1935 | 6-25 | 1.11 | 0.74 | 1.68 | Multivariabled |
| CGLF Study | Li et al., 2012 | Born between 1930-1964 vs. 1965-1999 | Stomach cancer mortality | >36 (24-month peak) | - | - | - | 63 | 26 | (see exposure contrast) | In utero-29 | 2.39 | 1.51 | 3.77 | Poisson (age, birth cohorts, and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 377,733 | 40,244 | 263 | 161 | Birth Cohort 1935-1939 | In utero-10 | 1.66 | 1.04 | 2.65 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 387,166 | 112,007 | 452 | 227 | Birth Cohort 1930-1934 | 5-15 | 2.04 | 1.59 | 2.61 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 492,743 | 132,929 | 755 | 355 | Birth Cohort 1925-1929 | 10-20 | 2.27 | 1.89 | 2.72 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 519,343 | 161,125 | 944 | 606 | Birth Cohort 1920-1924 | 15-25 | 1.59 | 1.39 | 1.83 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | | | | | Pooled estimateb | In utero-25 | 1.89 | 1.56 | 2.29 | |

The table Continues on the next page

S7 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Respiratory cancer mortality | 28 | 300 (winter 1941-42) | | | 87 | 117 | Birth Cohort 1916-1935 | 6-25 | 1.31 | 0.99 | 1.74 | Multivariabled |
| Females | | | | | | | | | | | | | | | |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 377,733 | 40,244 | 374 | 197 | Birth Cohort 1935-1939 | In utero-10 | 1.93 | 1.25 | 3.00 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 387,166 | 112,007 | 638 | 427 | Birth Cohort 1930-1934 | 5-15 | 1.51 | 1.25 | 1.82 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 492,743 | 132,929 | 1,151 | 807 | Birth Cohort 1925-1929 | 10-20 | 1.52 | 1.31 | 1.75 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | 519,343 | 161,125 | 1,617 | 1,305 | Birth Cohort 1920-1924 | 15-25 | 1.33 | 1.19 | 1.48 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Colorectal cancer incidence | 72 | 220-800 | | | | | Pooled estimateb | | 1.45c | 1.31 | 1.61 | |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Colorectal cancer mortality | 28 | 300 (winter 1941-42) | | | 9 | 13 | Birth Cohort 1916-1935 | 6-25 | 0.65c | 0.27 | 1.56 | Multivariabled |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 8,117 | 16,070 | 327 | 658 | Birth Cohort 1916-1932 | 12-28 | 0.99c | 0.83 | 1.18 | Multivariablee |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 232 | 390 | 2 | 15 | - | Before growth spurt | Not estimable (too few cases) | | | |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 1,840 | 3,511 | 41 | 88 | - | During growth spurt (2 yr< menarche<1 yr) | 0.89 | 0.57 | 1.39 | Multivariablee |
| NLCS | Hughes et al. 2010 | Lived in Western city vs. non-Western area | Colorectal cancer incidence | 5 | <700 (from January 1945 onwards) | 6,045 | 12,169 | 277 | 538 | - | After growth spurt | 1.08 | 0.89 | 1.32 | Multivariablee |

The table Continues on the next page

S7 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Stomach cancer mortality | 28 | 300 (winter 1941-42) | | | 18 | 16 | Birth Cohort 1916-1935 | 6-25 | 0.47 | 0.20 | 1.12 | Multivariabled |
| CGLF Study | Li et al., 2012 | Born between 1930-1964 vs. 1965-1999 | Stomach cancer mortality | >36 (24-month peak) | - | - | - | 45 | 28 | (see exposure contrast) | In utero-29 | 1.64 | 1.02 | 2.62 | Poisson (age, birth cohorts, and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 377,733 | 40,244 | 121 | 117 | Birth Cohort 1935-1939 | In utero-10 | 1.05 | 0.56 | 1.95 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 387,166 | 112,007 | 261 | 136 | Birth Cohort 1930-1934 | 5-15 | 1.93 | 1.39 | 2.68 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 492,743 | 132,929 | 447 | 357 | Birth Cohort 1925-1929 | 10-20 | 1.35 | 1.09 | 1.69 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | 519,343 | 161,125 | 545 | 493 | Birth Cohort 1920-1924 | 15-25 | 1.15 | 0.96 | 1.37 | Poisson (age and period) |
| Israeli Cohort | Keinan-Boker et al., 2009 | Immigrated after the war vs. before the war | Lung and Bronchial cancer incidence | 72 | 220-800 | | | | | Pooled estimateb | | 1.35 | 1.06 | 1.72 | |
| US-USSR LP | Koupil et al., 2009 | Lived in Leningrad vs. outside Leningrad | Respiratory cancer mortality | 28 | 300 (winter 1941-42) | - | - | 6 | 6 | Birth Cohort 1916-1935 | 6-25 | 0.98 | 0.31 | 3.15 | Multivariabled |
| NLCS | Schouten et al., 2012 | Lived in Western city vs. non-Western area | Ovarian cancer incidence | 5 | <700 (from January 1945 onwards) | 6,790 | 13,248 | 66 | 157 | Birth Cohort 1916-1931 | 13-28 | 0.97 | 0.68 | 1.39 | Multivariablef |
| NLCS | Schouten et al., 2012 | Lived in Western city vs. non-Western area | Ovarian cancer incidence | 5 | <700 (from January 1945 onwards) | 3,413 | 7,346 | 36 | 71 | Birth Cohort 1926-1931 | 13-19 | 1.11 | 0.71 | 1.72 | Multivariablef |

The table Continues on the next page

S7 Table: Continued

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NLCS | Schouten et al., 2012 | Lived in Western city vs. non-Western area | Ovarian cancer incidence | 5 | <700 (from January 1945 onwards) | 2,910 | 6,122 | 26 | 82 | Birth Cohort 1921-1926 | 18-24 | 0.64 | 0.40 | 1.04 | Multivariablef |
| NLCS | Schouten et al., 2012 | Lived in Western city vs. non-Western area | Ovarian cancer incidence | 5 | <700 (from January 1945 onwards) | 2,684 | 4,977 | 37 | 58 | Birth Cohort 1916-1921 | 23-29 | 1.12 | 0.72 | 1.76 | Multivariablef |
| Males and females | | | | | | | | | | | | | | | |
| NLCS | Heinen et al., 2011 | Lived in Western city vs. non-Western area | Pancreatic cancer incidence | 5 | <700 (from January 1945 onwards) | 11,821 | 25,058 | 75 | 176 | Birth cohort 1916-1932 | 12-28 | 0.89 | 0.67 | 1.19 | Multivariableg |

Abbreviations: CGL, Chinese Great Leap Forward; CI, confidence interval; NLCS, Netherlands Cohort Study; PY, person-years; US-USSR LP, US-USSR Lipid Program.

a In order to facilitate comparison between studies, birth cohort was derived from age at exposure if not mentioned in the paper and vice versa, and may therefore be an approximation.

a Even though heterogeneity in estimates between birth cohorts was >50% for some cancer endpoints studied by Keinan-Boker et al., all birth cohorts were used in the pooling, because no specific birth cohort introduced heterogeneity across endpoints as based on the leave-one-out estimates of heterogeneity and because the pooled estimate did not materially differ when excluding the specific birth cohort introducing heterogeneity in analyses for a certain cancer endpoint.

c Relative risk ratio or hazard ratio taken forward to the meta-analysis. If a cohort reported relative risk ratios or hazard ratios for multiple birth cohorts, a pooled estimate was calculated for these birth cohorts and the pooled estimate was taken forward to the meta-analysis.

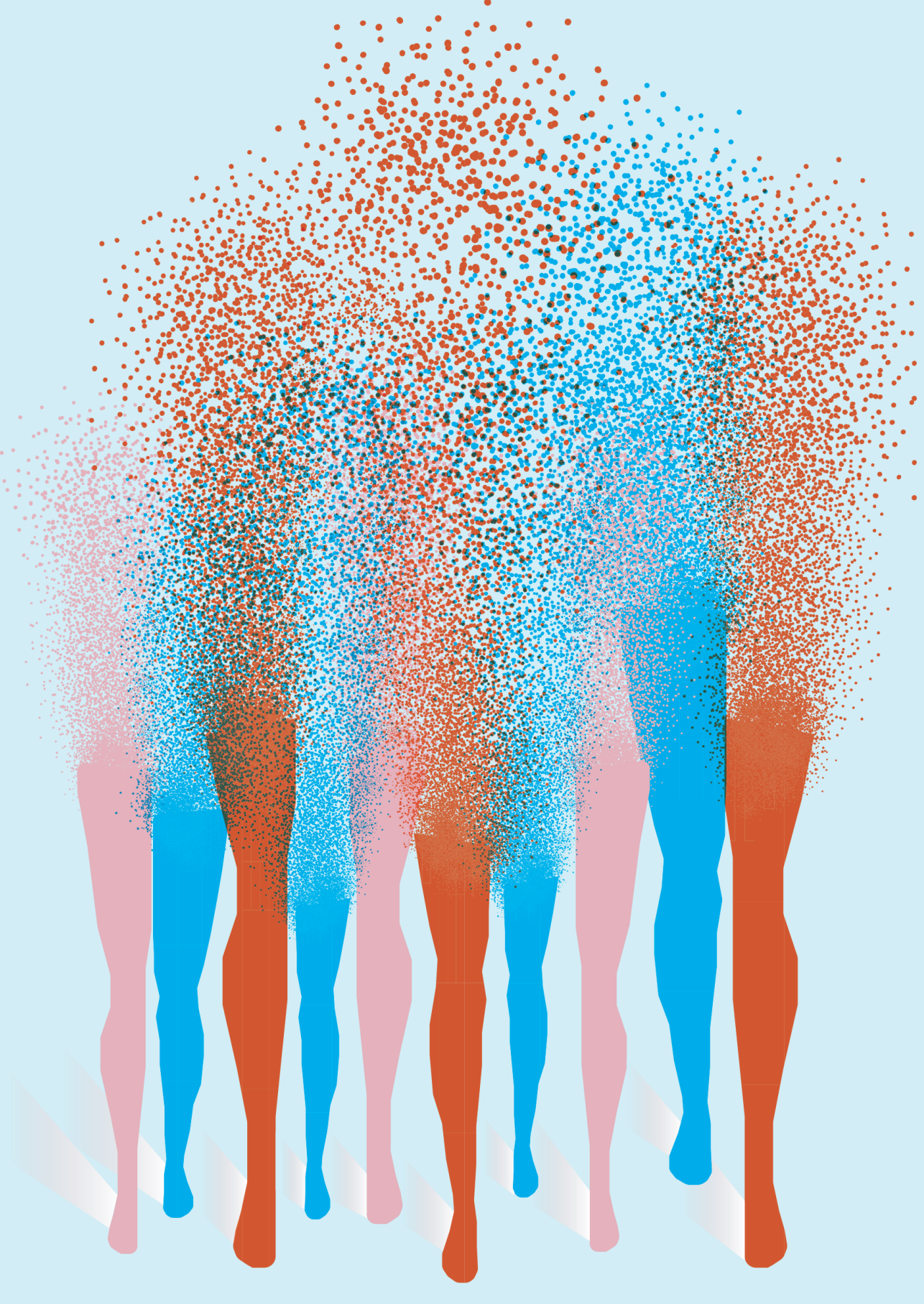
d Model adjusted for exact age at examination and social characteristics (education, occupation, ethnic group/nationality and marital status), behavioral characteristics (smoking, alcohol consumption) and adult stature (adult height and body mass index).

e Model adjusted for age, family history of colorectal cancer, body mass index (kg/m2), energy intake at baseline (kcal/day), alcohol intake at baseline (g/day), recreational physical activity (min/day), level of education, and smoking status (yes, ex, never).

f Model adjusted for age, number of children (continuous), use of oral contraceptives (never, ever), for hysterectomy (no, possible/probable).

g Model adjusted for age, sex, smoking, body mass index (kg/m2), baseline nonoccupational physical activity (min/d), history of sports participation (never or ever), energy intake (kcal/d) and intake of vegetables (g/d).

Supplemental table 7 continuous on the next page



Chapter 3

# Associations of Adult-Attained Height and Early Life Energy Restriction with Postmenopausal Breast Cancer Risk According to Estrogen and Progesterone Receptor Status

Rachel JJ Elands, Nadine SM Offermans, Colinda CJM Simons , Leo J Schouten, Bas AJ Verhage, Piet A van den Brandt, Matty P Weijenberg – *Under revision at the International Journal of Cancer*

## Abstract

**Background:** Adult-attained height is a marker for underlying mechanisms, such as cell growth, that may also influence postmenopausal breast cancer (BC) risk, perhaps specifically hormone-sensitive BC subtypes. Early life energy restriction may inhibit these mechanisms, resulting in shorter height and a reduced postmenopausal BC risk.

**Methods**: 62,573 women from the Netherlands Cohort Study, 55-69 years old, completed a self-administered questionnaire in 1986, and were followed-up for 20.3 years (case-cohort: $N_{subcohort}$=2438; $N_{cases}$=3354). Cox multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) were estimated for BC risk overall and by estrogen and progesterone receptor subtypes in relation to height and early life energy restriction during the Hunger Winter, War Years, and Economic Depression. Although energy restriction can only influence longitudinal growth in women exposed before and/or during the growth spurt, it may also influence BC risk when occurring after the growth spurt, possibly through different growth processes. Therefore, Cox analyses were additionally conducted according to timing of energy restriction in relation to the growth spurt.

**Results**: Height was associated with an increased BC risk ($HR_{per\ 5cm}$=1.07, 95%CI:1.01–1.13), particularly hormone receptor-positive BC. Energy restriction before and/or during the growth spurt was associated with a reduced hormone receptor-positive BC risk. Energy restriction during the Hunger Winter increased the estrogen receptor-negative BC risk regardless of the timing of energy restriction.

**Conclusions**: Height and energy restriction before and/or during the growth spurt were both associated with hormone receptor-positive BC risk, in the direction as expected, indicating critical exposure windows for hormonal growth-related mechanisms.

## Introduction

Based on meta-analyses from cohort studies, there is convincing evidence that postmenopausal breast cancer (BC) risk is increased by 7-11% for every 5 cm increase in adult-attained height.[1-3] This may be of particular relevance for the tallest population in the world, the population of the Netherlands,[4] which also has one of the highest BC rates worldwide.[5] While information on BC hormone receptor status is not always available, associations between adult-attained height and postmenopausal BC risk may exist particularly for the most commonly diagnosed breast cancer subtype, which is the estrogen receptor positive (ER+) subtype. Associations of adult-attained height with ER+ BC have been reported both in combination with PR+ BC[6,7] and when studied separately as an endpoint (*i.e.*, no information on PR status was available).[8-10] Significant increased ER+ BC relative risks have been reported in association with adult-attained height and null associations have been reported for ER- BC risk,[8,9] when ER status is studied regardless of PR status as an endpoint. Joint ER+PR+ BC status has been associated with increased relative risk in relation to adult-attained height,[6,7] whereas joint ER-PR- BC status has been associated both with increased relative risks[7] and null associations.[6] There is only one study that investigated adult-attained height in relation to PR+ tumors as a separate endpoint, which reported associations with non-significant increased relative risks for both PR+ and PR- BC status.[8] The inconclusive findings for adult-attained height and hormone-receptor negative subtypes, *i.e.* both null associations and non-significant increased relative risks, may be due to the fact that BCs with a negative hormone receptor status are less common, resulting in relatively less power to investigate these subtypes, and that information on PR BC status is less often available than ER BC status, limiting opportunities to investigate this. Yet, it is also plausible that there is a null association between adult-attained height and hormone receptor-negative BC.

Adult-attained height is a marker for mechanisms such as cell growth, which determine both adult-attained height and postmenopausal BC risk. These growth processes can be influenced by early life environmental exposures.[11,12] Therefore, when energy balance is disturbed in early life, *e.g.* by energy restriction, this can have an effect on adult-attained height and postmenopausal BC risk later in life. In a recent meta-analysis of observational studies by our group, severe transient energy restriction during early life was associated with a 28% increased BC risk, though some of the underlying studies showed null results.[13] Conversely,

moderate energy restriction during early life for a longer period of time, as studied in animal experimental models and human ecological studies, was generally inversely associated with BC risk.[13] Biologically, the latter seems plausible as continuous moderate early life energy restriction (pre- and peripubertal) may lead to decreased growth factor levels,[14] which in turn may result in a shorter stature[15-17] and a reduced postmenopausal BC risk.[18] Earlier and more rapid childhood and pubertal growth, for instance during catch-up growth, on the other hand, appear to increase postmenopausal BC risk.[19-22] In our meta-analysis, results from a sub analysis on energy restriction in women aged 10-20 years indicated a 21% increased BC risk compared to women not exposed during that age period. A comparison with the summary risk estimate in women aged 0-10 years was not possible, however, as this estimate could not be computed due to a high between-study heterogeneity.[13] One source of heterogeneity in results may be differences in distribution of hormone receptor subtypes among BC cases. To our knowledge, only one report has been published on the association between energy restriction and postmenopausal BC risk by ER+/- and PR+/- status, investigating combinations of ER+/- and PR+/- subtypes.[23] Results of this study showed that women exposed to the Chinese famine, particularly those exposed after birth (0-3 years), had an increased risk of ER-PR-BC, ER-PR+BC, and ER+PR-BC, while no association was observed with ER+PR+BC.

The Netherlands Cohort Study (NLCS), a prospective cohort study that includes 58,279 men and 62,573 women, has data available on height, early life energy restriction, ER+/- and PR+/- status of breast cancer cases during follow-up and other covariates. The long follow-up of 20.3 years enabled us to study the following aims with sufficient power. Firstly, we investigated the association of adult-attained height with overall postmenopausal BC risk and by ER+/- and PR+/- status of the tumor. Secondly, we examined the association of early life energy restriction with overall postmenopausal BC risk and by ER+/- and PR+/- status of the tumor. Considering that different growth processes may be at play during different periods in life, we additionally conducted these analyses according to timing of energy restriction in relation to the growth spurt. Energy restriction was hypothesized to be able to affect longitudinal growth only when exposure occurred before and/or during the growth spurt, but an effect on BC risk could also exist for exposure after the growth spurt, and not necessarily through a relationship between energy restriction and height. The association between adult-attained height and early life energy restriction with regard to breast cancer risk has been studied in the NLCS previously. A positive association between

adult-attained height and breast cancer risk among postmenopausal women was observed, after 4.3 years of follow-up.[24] With regard to energy restriction, neither exposure to early life energy restriction regardless of timing of the growth spurt and exposure to energy restriction during the adolescent growth spurt were associated with breast cancer risk after 6.3 years of follow-up.[25]

# METHODS

## Study population and design

The Netherlands Cohort Study (NLCS) includes 58,279 men and 62,573 women, who were 55 to 69 years old at baseline in September 1986.[26] Participants completed a self-administered questionnaire at baseline on cancer risk factors. For efficiency reasons, the NLCS uses a case-cohort design in which cases are enumerated from the entire cohort and the person-time at risk is estimated from a subcohort. This subcohort, consisting of 5,000 men and women, was randomly selected immediately after baseline and independent of exposure. The follow-up for vital status and migration is performed through linkage to the Central Bureau of Genealogy and the municipal population registries (~100% completeness).[27][29] Cancer follow-up is performed through linkage to the population-based cancer registry and PALGA (Netherlands pathology database; >95% completeness).[28] Participants who reported a history of cancer at baseline (except skin cancer) were excluded leaving 4,774 subcohort members, of which 2,438 female subcohort members. A total of 3,354 postmenopausal BC cases were identified in the total cohort after 20.3 years of follow-up (September 17, 1986, until January 1, 2007) (ICD-O-1749). ER and PR receptor status was available for 59% and 47% of the cases, respectively, with 1,620 ER+ cases, 364 ER- cases, 1,009 PR+ cases, and 556 PR- cases (see **table 1** for an overview of the overlap between hormone receptor subtypes). If dietary data of participants were incomplete or inconsistent, these participants were excluded, leaving 2,248 female subcohort members and 3,094 postmenopausal BC cases, among which 1,507 ER+ cases, 334 ER- cases, 934 PR+ cases, and 519 PR- cases (see flow chart in **Figure 1**).

**Table 1. Overview of the overlap between hormone receptor subtypes of breast cancer in postmenopausal women of the Netherlands Cohort Study**

| | PR+ BC cases | PR- BC cases | BC cases with unknown PR status | *Total number of BC cases by ER status* |
|---|---|---|---|---|
| | N (%) | N (%) | N (%) | |
| ER+ BC cases | 979 (97,0%) | 308 (55,4%) | 333 (18,6%) | 1,620 |
| ER- BC cases | 30 (3,0%) | 245 (44,1%) | 89 (5,0%) | 364 |
| BC cases with unknown ER status | 0 (0,0%) | 3 (0,5%) | 1,367 (76,4%) | 1,370 |
| *Total number of BC cases by PR status* | 1,009 | 556 | 1,789 | 3,354 |

Abbreviation: BC, breast cancer; ER+, estrogen receptor positive; ER-, estrogen receptor negative; PR+, progesterone receptor positive; PR-, progesterone receptor negative

## Exposure assessment

Adult-attained height (cm) was self-reported on the baseline questionnaire. Height was defined as a continuous variable (per 5 cm increase) as well as a categorical variable in tertiles based on the distribution in the female subcohort. Early life energy restriction was measured through three proxy variables, as individual food intake data in early life of the cohort members were not available. The proxies covered three time periods in the Netherlands during which a part of the population experienced energy restriction, *i.e.* the Hunger Winter (the winter of 1944-45), the War Years (1940-44), and the Economic Depression (1932-40). For the Hunger Winter, which was at its height between December 1944 and May 1945, place of residence was used to approximate exposure to energy restriction as individuals living in a western city, and to a lesser extent a western rural area, were exposed to (severe) energy restriction. During this time period, official daily rations per capita were between 400-800 kilocalories for women living in a western city.[29,30] In a follow-up study in the NLCS, female subcohort members were asked if they really had experienced hunger during the Hunger Winter. Of the women who reported severe hunger, 80% lived in a western city during this winter.[25] These results indicate that place of residence during the Hunger Winter is an adequate proxy for exposure to energy restriction. Place of residence was based on the reported residence during the winter of 1944-45 and classified into 'living in a non-western area, 'living in a western rural area', or 'living in a western city'. Also for the War Years, place of residence was used as a proxy for energy restriction. During the War Years, food rationing was introduced, and caloric intake was reduced to ~1700 kcal/d during 1941–1943.[31] Between 1943 and 1944, the nutritional status of the Dutch population deteriorated, especially

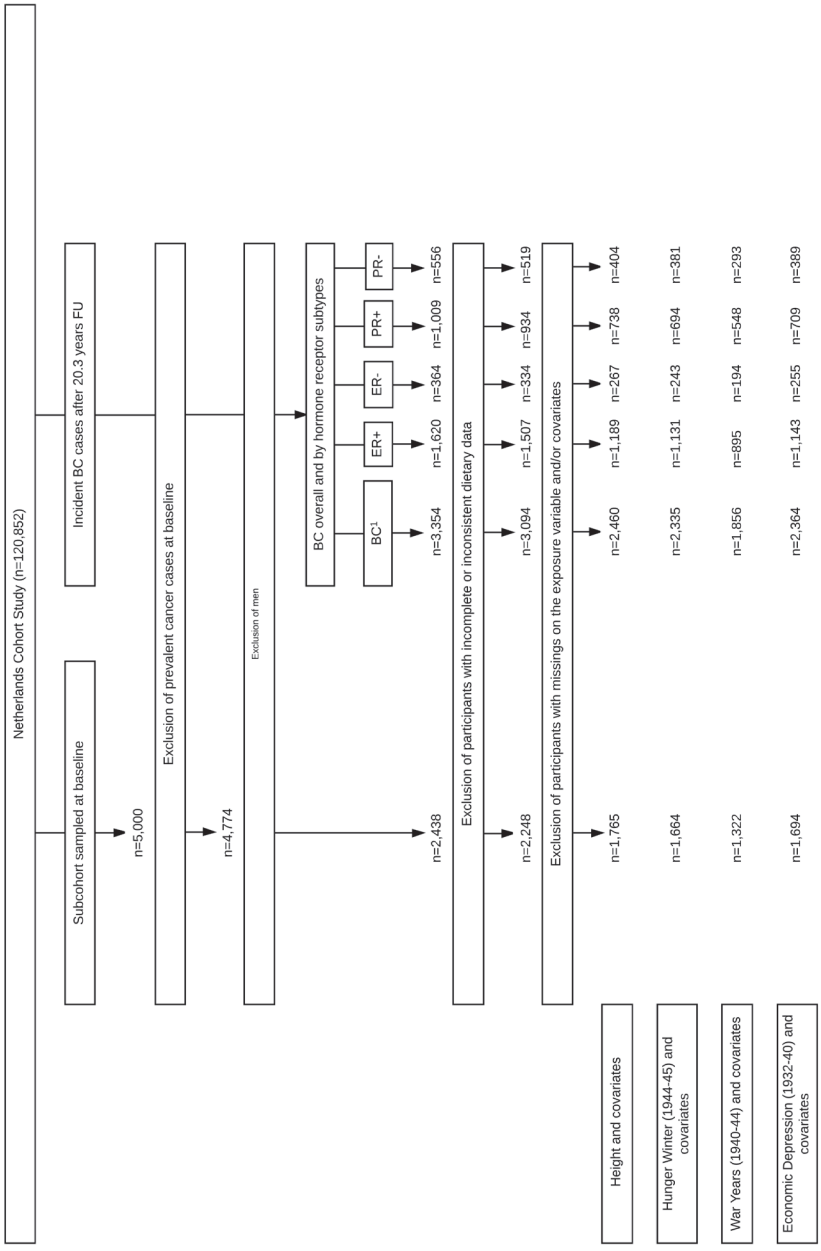



Figure 1. Flow diagram of available subcohort members and breast cancer cases among postmenopausal women in the Netherlands Cohort Study, 1986-2006

**Abbreviations: BC, breast cancer; FU, follow-up; ER+, estrogen receptor positive; ER-, estrogen receptor negative; PR+, progesterone receptor positive; PR-, progesterone receptor negative**

[1] The sum of cases of the different BC subtypes differs from the total number of BC cases as hormone receptor status was not known for all BC cases.

for those living in the cities.[25] Place of residence was based on the question to list the last 4 residences before baseline of the study, which resulted in a classification into 'living in an urban area' (defined as a town with at least 40,000 residents) or 'living in a rural area' in 1942 (the midpoint of the War Years 1940-44). For the Economic Depression, employment status of the father was used to approximate exposure to energy restriction. Several surveys showed that having an unemployed father indicated that the number of calories available was less and the variation in the individual's food pattern was limited compared with families with an employed father.[25,32] Father's employment status was dichotomized into participants whose father had a job during the years of the Economic Depression or worked intermittingly ('employed') and cohort members with fathers without a job during these years ('unemployed').[33] Information on covariates that were considered potential confounders on the basis of previous research was also available from the baseline questionnaire.

## Statistical Analyses

We calculated the mean adult-attained height for each category of the three proxies for early life energy restriction in those female subcohort members exposed to energy restriction before and/or during the growth spurt, since we only expected an effect of energy restriction on mean adult-attained height in this subgroup of women. The historic events that are associated to the three proxies took place in three different periods in time and thus affected cohort members at different ages (12-28 years for the Hunger Winter, 7-28 years for the War Years, and 0-23 years for the Economic Depression). The growth spurt was defined as two years before the self-reported age at menarche until one year after the reported age at menarche. Furthermore, in a sensitivity analysis, we restricted the analyses for the Hunger Winter and the War Years to individuals from the middle provinces of the Netherlands. We did so, because there was virtually no exposure to energy restriction in the northern and southern provinces during both time periods, as these are mostly non-western and rural areas, and, at the same time, these provinces are characterized by, on average, the tallest and shortest people in the country, respectively.[4] Including individuals from these provinces in the analyses, therefore, can mask the relationship between energy restriction and height, if present. Cox proportional hazards analysis was used to estimate both age-adjusted and multivariable-adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) for postmenopausal BC risk overall and by ER+/- and PR+/- status. In the multivariable-adjusted models, we adjusted for potential confounders that were selected *a priori* on the basis of that these were risk factors in the literature. To account for the additional variance introduced by sampling the

subcohort from the entire cohort, standard errors were estimated using the robust Huber-White sandwich estimator.[34] The proportional hazards assumption was tested using the scaled Schoenfeld residuals and by visual inspection of the -log (-log) transformed hazard curves.[35]

In the multivariable-adjusted models for height and postmenopausal BC risk, the following *a priori* selected potential confounders were included on the basis of that these are potential risk factors for postmenopausal BC: age at baseline (y), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency (numbers of cigarettes per day; continuous, centered), and smoking duration (number of years; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menarche (y), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever), and postmenopausal hormone-replacement therapy (never, ever). After excluding participants without (complete) information on height and/or the covariates, 1,765 subcohort members and 2,460 postmenopausal BC cases were left for analysis (see Figure 1).

Considering that (most) other studies that have investigated the height-BC association have not adjusted for body mass index (BMI) (kg/m$^2$) or weight (kg), our primary model will not include BMI or weight, enabling comparison of our results with those in the literature. BMI approximates body fatness, though may do so differentially depending on age and height, with a positive correlation between height and BMI in young populations and an inverse correlation in older populations.[36,37] Since the NLCS comprises an older population, we conducted two additional sensitivity analyses to investigate whether including either BMI or weight as additional continuous covariates in the models for height changed the associations.

In the multivariable-adjusted models for early life energy restriction and postmenopausal BC risk, the following *a priori* selected potential confounders were included on the basis of that these are potential risk factors for postmenopausal BC: age at baseline (y), BMI (kg/m$^2$), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency (numbers of

cigarettes per day; continuous, centered), smoking duration (number of years; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever), and postmenopausal hormone-replacement therapy (never, ever). In a sensitivity analysis, we checked whether additional adjustment for adult-attained height and age at menarche changed the results because these are also important risk factors for postmenopausal BC. However, early life energy restriction may also influence adult-attained height and age at menarche and these could thus also act as intermediate factors. After excluding participants without (complete) information on energy restriction and/or the primary covariates of interest, 1,664 subcohort members and 2,335 postmenopausal BC cases were left for analysis for the Hunger Winter, 1,322 subcohort members and 1,856 postmenopausal BC cases left for analysis for the War Years, and 1,694 subcohort members and 2,364 postmenopausal BC cases for the Economic Depression (see **Figure 1**).

# RESULTS

**Table 2** shows baseline characteristics of the female subcohort members and postmenopausal BC cases overall and by hormone receptor subtypes. Postmenopausal BC cases, in particular ER+ and PR+ cases, more often reported a family history of BC compared to subcohort members. Additionally, postmenopausal BC cases more often reported a history of benign breast disease compared to subcohort members.

**Table 2. Baseline characteristics of female subcohort members and postmenopausal BC cases overall and by hormone receptor subtypes in the Netherlands Cohort Study[a]**

| | Subcohort | | Postmenopausal BC cases | | ER+ | | ER- | | PR+ | | PR- | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean (SD) | (%) | Mean (SD) | (%) | Mean (SD) | (%) | Mean (SD) | (%) | Mean (SD) | (%) | Mean (SD) | (%) |
| Height, cm | 165.2 (6.2) | | 165.8 (6.5) | | 165.9 (6.5) | | 165.3 (6.7) | | 165.8 (6.4) | | 165.7 (6.6) | |
| Residence during Hunger Winter (1944-45)[b] | | | | | | | | | | | | |
| Non-western | | 56.6 | | 55.1 | | 55.6 | | 47.7 | | 62.4 | | 57.7 |
| Western rural | | 15.0 | | 16.3 | | 15.4 | | 21.2 | | 13.9 | | 15.0 |
| Western city | | 28.4 | | 28.6 | | 29.0 | | 31.1 | | 23.7 | | 27.3 |
| Residence during War years (1940-44)[b] | | | | | | | | | | | | |
| Rural area in 1942 | | 46.8 | | 47.1 | | 43.8 | | 44.8 | | 47.2 | | 44.3 |
| Urban area in 1942 | | 53.2 | | 52.9 | | 56.2 | | 55.2 | | 52.8 | | 55.7 |
| Job status father during Economic Depression (1932-40)[b] | | | | | | | | | | | | |
| Employed | | 88.5 | | 90.2 | | 90.1 | | 87.9 | | 90.8 | | 90.6 |
| Unemployed | | 11.5 | | 9.8 | | 9.9 | | 12.1 | | 9.2 | | 9.4 |
| Age at baseline, y | 61.4 (4.3) | | 61.3 (4.1) | | 61.3 (4.2) | | 61.0 (3.9) | | 61.3 (4.2) | | 60.9 (3.9) | |
| Adult BMI, kg/m$^2$ | 25.1 (3.6) | | 25.4 (3.4) | | 25.4 (3.3) | | 24.9 (3.4) | | 25.5 (3.4) | | 25.1 (3.3) | |
| Energy intake, kcal/d | 1686 (397) | | 1688 (399) | | 1683 (398) | | 1691 (416) | | 1681 (397) | | 1729 (413) | |
| Baseline non-occupational physical activity, min/d[b] | | | | | | | | | | | | |
| ≤30 | | 25.0 | | 27.9 | | 26.9 | | 28.6 | | 27.8 | | 26.9 |
| 30–≤60 | | 31.2 | | 31.8 | | 32.0 | | 24.9 | | 31.5 | | 30.4 |
| >60–90 | | 22.4 | | 21.1 | | 21.0 | | 24.6 | | 20.5 | | 22.0 |
| >90 | | 21.4 | | 19.2 | | 20.1 | | 21.9 | | 20.2 | | 20.7 |
| Cigarette smoking[b] | | | | | | | | | | | | |
| Never | | 58.4 | | 56.4 | | 56.1 | | 53.6 | | 57.2 | | 54.7 |
| Former | | 20.6 | | 22.6 | | 22.7 | | 23.0 | | 23.8 | | 20.6 |
| Current | | 21.0 | | 21.0 | | 21.2 | | 23.4 | | 19.0 | | 24.7 |

The table Continues on the next page

**Table 2. continued**

| | | | | | | |
|---|---|---|---|---|---|---|
| Number of cigarettes per day[c] | 4.6 (7.7) | 5.0 (8.0) | 5.0 (8.1) | 5.6 (8.4) | 4.6 (7.7) | 5.3 (8.3) |
| Years of smoking[c] | 11.4 (15.8) | 11.8 (15.9) | 12.0 (16.1) | 12.7 (16.2) | 11.4 (15.7) | 12.7 (16.4) |
| Alcohol intake, g/d[b] | | | | | | |
| 0 | 32.3 | 30.3 | 30.6 | 31.6 | 31.8 | 28.8 |
| 0.1–29 | 64.2 | 64.7 | 65.2 | 63.5 | 64.5 | 65.8 |
| ≥30 | 3.5 | 5.0 | 4.2 | 4.9 | 3.7 | 5.4 |
| Level of education[b] | | | | | | |
| Primary school | 33.5 | 32.5 | 31.7 | 35.0 | 31.5 | 35.1 |
| Lower vocational school | 23.2 | 21.6 | 22.2 | 19.0 | 21.5 | 19.4 |
| Intermediate vocational/high school | 34.5 | 36.8 | 37.6 | 36.6 | 37.7 | 37.7 |
| Higher vocational school/ university | 8.8 | 9.1 | 8.5 | 9.4 | 9.3 | 7.8 |
| Family history of BC[b] | | | | | | |
| No | 91.3 | 85.8 | 85.9 | 88.6 | 84.7 | 89.2 |
| Yes | 8.7 | 14.2 | 14.1 | 11.4 | 15.3 | 10.8 |
| History of benign breast disease[b] | | | | | | |
| No | 92.3 | 87.5 | 88.8 | 88.0 | 88.8 | 88.2 |
| Yes | 7.7 | 12.5 | 11.2 | 12.0 | 11.2 | 11.8 |
| Age at menarche, y | 13.7 (1.8) | 13.5 (1.7) | 13.5 (1.8) | 13.4 (1.7) | 13.4 (1.7) | 13.6 (1.8) |
| Age at menopause, y | 48.7 (4.5) | 49.1 (4.3) | 49.0 (4.4) | 48.6 (4.3) | 48.8 (4.5) | 49.1 (4.2) |
| Age at first childbirth[b], y | | | | | | |
| Nulliparous | 18.0 | 19.4 | 18.7 | 17.7 | 18.2 | 17.5 |
| <25 | 22.9 | 21.3 | 20.8 | 24.8 | 21.4 | 21.2 |
| ≥25 | 59.1 | 59.3 | 60.5 | 57.5 | 60.4 | 61.3 |
| Parity[b], *n* children | | | | | | |
| Nulliparous | 18.0 | 19.4 | 18.7 | 17.7 | 18.2 | 17.5 |
| 1 child | 8.3 | 8.9 | 9.2 | 8.7 | 9.8 | 7.3 |
| 2 children | 21.2 | 23.7 | 24.0 | 22.1 | 23.1 | 23.3 |
| ≥3 children | 52.5 | 48.0 | 48.1 | 51.5 | 48.9 | 51.9 |

The table Continues on the next page

**Table 2. continued**

| | | | | | | |
|---|---|---|---|---|---|---|
| Oral contraceptive use[b] | | | | | | |
| Never | 75.0 | 75.0 | 76.1 | 72.2 | 77.7 | 72.2 |
| Ever | 25.0 | 25.0 | 23.9 | 7.8 | 22.3 | 27.8 |
| Postmenopausal hormone replacement therapy[b] | | | | | | |
| No | 86.8 | 86.4 | 86.0 | 87.7 | 84.8 | 86.6 |
| Yes | 13.2 | 13.6 | 14.0 | 12.3 | 15.2 | 13.4 |

Abbreviations: BC, breast cancer; BMI, body mass index; ER+, estrogen receptor positive; ER-, estrogen receptor negative; PR+, progesterone receptor positive; PR-, progesterone receptor negative; SD, standard deviation.

[a] Participants with incomplete or inconsistent dietary data were excluded.

[b] Sums of categories differ because of missing values on the particular variable.

[c] Among former and current smokers only.

## Mean adult-attained height according to early life energy restriction

**Table 3** shows the mean adult-attained height by exposure to early life energy restriction in female subcohort members who were exposed to energy restriction before and/or during the growth spurt, and when further restricting to female subcohort members living in the middle provinces of the Netherlands. There were no differences in height between female subcohort members who were and those who were not exposed to energy restriction before and/or during the growth spurt as a result of the Hunger Winter or the War Years. Further restriction of the analyses to those living in the middle provinces of the Netherlands did not change this finding. During the Economic Depression, most of the women from the cohort were younger (0-23 years) and most (94.4 %) before or in their growth spurt their growth spurt. Female subcohort members who were exposed to energy restriction during the Economic Depression before or during their growth spurt were statistically significantly shorter than those without this exposure (163.8 cm versus 165.5 cm, respectively).

## Adult-attained height and postmenopausal BC risk

**Table 4** shows the associations between adult-attained height and postmenopausal BC risk overall and by hormone receptor subtypes. Height was associated with a significantly increased risk of postmenopausal BC overall ($HR_{per\ 5\ cm}$=1.07, 95% CI: 1.01–1.13). In particular, height was positively associated with ER+ BC risk ($HR_{per\ 5\ cm}$=1.08, 95% CI: 1.01–1.15) and, though borderline significant, with PR+ BC risk ($HR_{per\ 5\ cm}$=1.07, 95% CI: 0.99–1.16), but not with ER- BC risk ($HR_{per\ 5\ cm}$: 1.03, 95% CI: 0.92-1.16) or PR- BC risk ($HR_{per\ 5\ cm}$=1.05, 95% CI: 0.95–1.16). Including either BMI or weight, both as additional continuous covariates, in the models for height did not change the associations (data not shown).

**Supplemental table 1** shows the associations between adult-attained height and postmenopausal BC risk by combinations of hormone receptor subtypes. Height was only borderline significantly associated with an increased ER+PR- BC risk ($HR_{per\ 5\ cm}$=1.07, 95% CI: 0.99–1.16) and an increased ER+PR- BC risk ($HR_{per\ 5\ cm}$=1.12, 95% CI: 0.99–1.28), though the number of cases was rather low for other combinations, especially ER-PR+ BC.

**Table 3. Mean and standard deviation of adult-attained height by exposure to early life energy restriction in all female subcohort members in the Netherlands Cohort Study, when restricting to those who were exposed to energy restriction before and/or during the growth spurt[a], and when additionally restricting to those living in the middle provinces of the Netherlands**

| Proxies for early life energy restriction | All women | | Women with energy restriction before/during the growth spurt | | Women with energy restriction before/during the growth spurt who lived in the middle provinces of the Netherlands | |
|---|---|---|---|---|---|---|
| | N | Mean height in cm (SD) | N | Mean height in cm (SD) | N | Mean height in cm (SD) |
| Residence during Hunger Winter (1944-45) | | | | | | |
| Non-western | 1,150 | 165.5 (6.3) | 206 | 165.9 (6.6) | 44 | 165.0 (6.7) |
| Western rural | 306 | 165.6 (6.0) | 57 | 164.9 (5.6) | 57 | 164.9 (5.6) |
| Western city | 583 | 165.4 (5.8) | 108 | 166.5 (5.7) | 108 | 166.5 (5.7) |
| *P*-value[b] | | 0.91 | | 0.32 | | 0.18 |
| Residence during War Years (1940-44) | | | | | | |
| Rural area in 1942 | 753 | 165.6 (6.0) | 267 | 165.4 (6.1) | 114 | 165.0 (5.7) |
| Urban area in 1942 | 858 | 165.1 (6.0) | 294 | 165.5 (5.9) | 201 | 165.9 (5.6) |
| *P*-value[b] | | 0.11 | | 0.79 | | 0.14 |
| Job status father during Economic Depression (1932-40) | | | | | | |
| Employed | 1,836 | 165.5 (6.1) | 1,736 | 165.5 (6.1) | -[c] | -[c] |
| Unemployed | 240 | 163.6 (6.8) | 222 | 163.8 (6.6) | -[c] | -[c] |
| *P*-value[b] | | <0.0001 | | 0.0001 | | |

Abbreviation: SD, standard deviation.

[a] The growth spurt was defined as: 2 years before reported age of menarche till 1 year after reported age of menarche.

[b] Testing for significant differences in mean height between categories of the proxies for early life energy restriction was performed using a t-test in case of two categories and ANOVA in case of more than two categories, with $p<0.05$ being significant.

[c] For the Economic Depression, the analyses were not restricted to individuals from the middle provinces of the Netherlands as exposure to energy restriction during this time period was not based on place of residence.

**Table 4. Hazard ratios and 95% confidence intervals for the association between adult-attained height and postmenopausal BC risk overall and by hormone receptor subtypes in the Netherlands Cohort Study, 1986-2006**

| | | Postmenopausal BC | | | ER+ | | | ER- | | | PR+ | | | PR- | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PY | N | HR[a] | 95%CI | N | HR[a] | 95%CI | N | HR[a] | 95%CI | N | HR[a] | 95%CI | N | HR[a] | 95%CI |
| Height (cm)[b] | | | | | | | | | | | | | | | | |
| Tertile 1 (range: 132-163 cm) | 11,424 | 830 | 1 | ref | 396 | 1 | ref | 105 | 1 | ref | 249 | 1 | | 147 | 1 | ref |
| Tertile 2 (range: 164-168 cm) | 10,609 | 864 | 1.12 | (0.96-1.31) | 424 | 1.15 | (0.96-1.39) | 86 | 0.88 | (0.64-1.22) | 266 | 1.15 | (0.93-1.44) | 130 | 0.94 | (0.72-1.23) |
| Tertile 3 (range: 169-198 cm) | 8,595 | 766 | 1.23 | (1.04-1.45) | 369 | 1.25 | (1.03-1.52) | 76 | 0.97 | (0.69-1.37) | 223 | 1.19 | (0.94-1.50) | 127 | 1.14 | (0.86-1.51) |
| *p* for trend | | | 0.02 | | | 0.02 | | | 0.81 | | | 0.13 | | | 0.41 | |
| Continuous per 5 cm | 30,629 | 2,460 | 1.07 | (1.01-1.13) | 1,189 | 1.08 | (1.01-1.15) | 267 | 1.03 | (0.92-1.16) | 738 | 1.07 | (0.99-1.16) | 404 | 1.05 | (0.95-1.16) |

Abbreviations: BC, breast cancer; CI, confidence interval; ER+, estrogen receptor positive; ER-, estrogen receptor negative; HR, hazard ratio; PR+, progesterone receptor positive; PR-, progesterone receptor negative; PY, person-years.

[a] Adjusted for age (y), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency and smoking duration (number of cigarettes per day and number of years, respectively; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menarche (y), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever),and postmenopausal hormone-replacement therapy (never, ever).

[b] Tertiles were based on the distribution in the subcohort.

## Early life energy restriction and postmenopausal BC risk

**Table 5** shows the associations between exposure to early life energy restriction and postmenopausal BC risk overall and by hormone receptor subtypes, including stratification on exposure to energy restriction before and/or during the growth spurt versus after the growth spurt. Exposure to energy restriction before and/or during the growth spurt could potentially have an effect on longitudinal growth and, particularly in this group; exposure to energy restriction was associated with a decreased risk of postmenopausal BC, particularly ER+ BC and PR+ BC. This result was observed across all three proxies. More specifically, women residing in a western city during the Hunger Winter compared to women residing in a non-western area had a significantly decreased risk of ER+ BC and PR+ BC (HR=0.49; 95% CI: 0.28–0.88; HR=0.23; 95% CI: 0.10–0.54, respectively). Women living in an urban area during the War Years compared to women living in a rural area had a (non)significantly decreased risk of ER+ BC and PR+ BC (HR=0.72; 95% CI: 0.51–1.01; HR=0.59; 95% CI: 0.39–0.89, respectively). Women with an unemployed father during the Economic Depression compared to women with an employed father had a non-significantly decreased risk of ER+ BC and PR+ BC (HR=0.89; 95% CI: 0.68–1.17; HR=0.76; 95% CI: 0.54–1.07, respectively). The risk of ER- BC was significantly increased for women residing in a western city during the Hunger Winter compared to women residing in a non-western area without stratification on timing of exposure to energy restriction in relation to women's growth spurt (HR=1.54; 95% CI: 1.11–2.12). This increased risk of ER- BC also seemed independent of whether exposure to energy restriction was before and/or during or after the growth spurt, as both hazard ratios were increased, although the first was not statistically significantly increased (HR=1.82; 95% CI: 0.69–4.78 and HR=1.51; 95% CI: 1.06–2.17, respectively). This non-significant finding for women exposed before and/or during the growth spurt may be due to a low number of cases in this subgroup. Except for the association of exposure to energy restriction during the Hunger Winter when exposed after the growth spurt with increased ER- BC risk, no significant associations were observed between exposure to energy restriction after the growth spurt with regard to the other proxies of energy restriction, *i.e.* the War Years and Economic Depression and other subtypes of BC. Additional adjustment for adult-attained height and age at menarche did not change these results (data not shown).

**Table 5. Hazard ratios and 95% confidence intervals for the association between exposure to early life energy restriction and postmenopausal BC risk overall and by hormone receptor subtypes, including stratification by timing of the growth spurt[a] in the Netherlands Cohort Study, 1986-2006**

| | Postmenopausal BC risk | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All women | | | | Women with energy restriction before/during the growth spurt | | | | Women with energy restriction after the growth spurt | | | |
| Proxies for energy restriction by cancer endpoint | PY | N | HR[b] | 95%CI | PY | N | HR[b] | 95%CI | PY | N | HR[b] | 95%CI |
| *Postmenopausal BC* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 1,280 | 1 | (ref) | 3,042 | 204 | 1 | (ref) | 13,507 | 1,076 | 1 | (ref) |
| Western rural | 4,150 | 377 | 1.18 | (0.97-1.44) | 791 | 65 | 1.50 | (0.90-2.51) | 3,359 | 312 | 1.14 | (0.92-1.42) |
| Western city | 8,250 | 678 | 1.03 | (0.88-1.20) | 1,603 | 101 | 0.68 | (0.43-1.08) | 6,647 | 577 | 1.09 | (0.91-1.29) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 859 | 1 | (ref) | 3,899 | 305 | 1 | (ref) | 6,641 | 554 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 997 | 0.93 | (0.79-1.10) | 4,705 | 313 | 0.72 | (0.54-0.97) | 7,681 | 684 | 1.06 | (0.87-1.31) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 2,146 | 1 | (ref) | 25,078 | 2,029 | 1 | (ref) | 1,169 | 117 | 1 | (ref) |
| Unemployed | 3,151 | 218 | 0.82 | (0.66-1.02) | 2,900 | 206 | 0.86 | (0.68-1.08) | 251 | 12 | 0.33 | (0.10-1.09) |
| *ER+* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 624 | 1 | (ref) | 3,042 | 109 | 1 | (ref) | 13,507 | 515 | 1 | (ref) |
| Western rural | 4,150 | 181 | 1.17 | (0.93-1.48) | 791 | 35 | 1.51 | (0.82-2.77) | 3,359 | 146 | 1.12 | (0.87-1.46) |
| Western city | 8,250 | 326 | 1.02 | (0.85-1.24) | 1,603 | 42 | 0.49 | (0.28-0.88) | 6,647 | 284 | 1.12 | (0.91-1.38) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 394 | 1 | (ref) | 3,899 | 154 | 1 | (ref) | 6,641 | 240 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 501 | 1.05 | (0.87-1.28) | 4,705 | 153 | 0.72 | (0.51-1.01) | 7,681 | 348 | 1.29 | (1.00-1.67) |

The table Continues on the next page

**Table 5. continued**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 1,037 | 1 | (ref) | 25,078 | 981 | 1 | (ref) | 1,169 | 56 | 1 | (ref) |
| Unemployed | 3,151 | 106 | 0.83 | (0.64-1.09) | 2,900 | 102 | 0.89 | (0.68-1.17) | 251 | 4 | –[c] | –[c] |
| *ER-* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 108 | 1 | (ref) | 3,042 | 20 | 1 | (ref) | 13,507 | 88 | 1 | (ref) |
| Western rural | 4,150 | 53 | 2.10 | (1.44-3.08) | 791 | 4 | –[c] | –[c] | 3,359 | 49 | 2.24 | (1.49-3.38) |
| Western city | 8,250 | 82 | 1.54 | (1.11-2.12) | 1,603 | 17 | 1.82 | (0.69-4.78) | 6,647 | 65 | 1.51 | (1.06-2.17) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 82 | 1 | (ref) | 3,899 | 25 | 1 | (ref) | 6,641 | 57 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 112 | 1.05 | (0.75-1.48) | 4,705 | 40 | 1.07 | (0.60-1.92) | 7,681 | 72 | 1.07 | (0.69-1.65) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 222 | 1 | (ref) | 25,078 | 212 | 1 | (ref) | 1,169 | 10 | 1 | (ref) |
| Unemployed | 3,151 | 33 | 1.23 | (0.81-1.86) | 2,900 | 31 | 1.24 | (0.81-1.92) | 251 | 2 | –[c] | –[c] |
| *PR+* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 431 | 1 | (ref) | 3,042 | 75 | 1 | (ref) | 13,507 | 356 | 1 | (ref) |
| Western rural | 4,150 | 96 | 0.88 | (0.66-1.17) | 791 | 17 | 1.04 | (0.49-2.23) | 3,359 | 79 | 0.85 | (0.62-1.17) |
| Western city | 8,250 | 167 | 0.74 | (0.59-0.93) | 1,603 | 17 | 0.23 | (0.10-0.54) | 6,647 | 150 | 0.83 | (0.65-1.07) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 257 | 1 | (ref) | 3,899 | 102 | 1 | (ref) | 6,641 | 155 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 291 | 0.93 | (0.73-1.18) | 4,705 | 88 | 0.59 | (0.39-0.89) | 7,681 | 203 | 1.16 | (0.85-1.58) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 652 | 1 | (ref) | 25,078 | 612 | 1 | (ref) | 1,169 | 40 | 1 | (ref) |
| Unemployed | 3,151 | 57 | 0.71 | (0.51-0.98) | 2,900 | 55 | 0.76 | (0.54-1.07) | 251 | 2 | –[c] | –[c] |
| *PR-* | | | | | | | | | | | | |

The table Continues on the next page

**Table 5. continued**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 213 | 1 | (ref) | 3,042 | 40 | 1 | (ref) | 13,507 | 173 | 1 | (ref) |
| Western rural | 4,150 | 61 | 1.22 | (0.88-1.70) | 791 | 12 | 1.68 | (0.69-4.07) | 3,359 | 49 | 1.16 | (0.80-1.69) |
| Western city | 8,250 | 107 | 1.02 | (0.77-1.34) | 1,603 | 23 | 1.05 | (0.51-2.19) | 6,647 | 84 | 1.00 | (0.74-1.37) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 125 | 1 | (ref) | 3,899 | 46 | 1 | (ref) | 6,641 | 79 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 168 | 1.08 | (0.80-1.44) | 4,705 | 59 | 0.90 | (0.55-1.48) | 7,681 | 109 | 1.25 | (0.85-1.82) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 353 | 1 | (ref) | 25,078 | 339 | 1 | (ref) | 1,169 | 14 | 1 | (ref) |
| Unemployed | 3,151 | 36 | 0.84 | (0.57-1.26) | 2,900 | 36 | 0.92 | (0.61-1.38) | 251 | 0 | –[c] | –[c] |

Abbreviations: BC, breast cancer; CI, confidence interval; ER+, estrogen receptor positive; ER-, estrogen receptor negative; HR, hazard ratio; PR+, progesterone receptor positive; PR-, progesterone receptor negative; PY, person-years.

[a] The growth spurt was defined as: 2 years before reported age of menarche till 1 year after reported age of menarche.

[b] Adjusted for age (y), BMI (kg/m$^2$), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency and smoking duration (number of cigarettes per day and number of years, respectively; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever), and postmenopausal hormone-replacement therapy (never, ever).

[c] Estimate was unstable because of a low number of cases.

**Supplemental table 2** shows the associations between exposure to early life energy restriction and postmenopausal BC risk by combinations of hormone receptor subtypes, including stratification on timing of exposure in relation to the growth spurt. Again, when restricting to those women who were exposed to energy restriction before and/or during the growth spurt, the period during which exposure to energy restriction could potentially have an effect on longitudinal growth, early life energy restriction was (borderline) significantly associated with a decreased risk of ER+PR+ BC, which was consistently observed across all three proxies (HR=0.23; 95% CI: 0.10–0.54 for women living in a western city during the Hunger Winter compared to women residing in a non-western area; HR=0.60; 95% CI: 0.39–0.91 for women living in an urban area during the War Years compared to women living in a rural area; and HR=0.74; 95% CI: 0.53–1.04 for women with an unemployed father during the Economic Depression compared to women with an employed father). The three remaining combinations of hormone receptor subtypes showed no significant associations with any of the energy restriction exposures considered. However, it should be kept in mind that the number of cases was rather low, especially for ER-PR+ BC and ER-PR- BC; resulting in unstable HRs and we refrained from presenting these when the number of cases in the exposure or reference category was less than five.



## DISCUSSION

In this study, height was significantly positively associated with BC risk, in particular with hormone receptor-positive BC subtypes. Of the three exposures to energy restriction investigated, *i.e.* exposure to energy restriction during the Economic Depression, War Years, and Hunger Winter, only exposure to energy restriction during the Economic Depression was related to a shorter stature in female subcohort members who were before and/or during their growth spurt, and thus relatively young (0-23 years). Nevertheless, energy restriction during all three periods of exposure provided it occurred before and/or during the growth spurt was associated with a significantly decreased risk of hormone receptor-positive BC subtypes. Interestingly, exposure to energy restriction during the Hunger Winter was also associated with an increased ER- BC risk, which seemed independent of timing of exposure to energy restriction in relation to women's growth spurt.

In agreement with previous studies,[1-3] we observed a 7% increased risk in postmenopausal BC per 5 cm increase in adult-attained height. Previously in the NLCs, after 4.3 years of follow-up, the association between adult-attained height and breast cancer risk has also been studied and a positive association was reported, however result were not stratified by hormone receptor-defined subtypes among postmenopausal women.[24] Regarding hormone receptor status, a recent meta-analysis reported that a positive association between adult-attained height and BC risk was primarily limited to hormone receptor-positive BC, both ER+ and PR+ BC separately as well as combined ER+ and PR+ status.[38] A borderline significant positive association was observed between adult-attained height and PR- BC.[38] While this meta-analysis did not distinguish between premenopausal and postmenopausal BC cases and ER status was only known for 8.7% of the cases and PR status for 5.4% of the cases (compared to 59% and 47% of the cases, respectively, in our study), it is in support of our finding that adult-attained height may be particularly associated with hormone receptor-positive BC subtypes. Such an association points to the involvement of hormone-related growth-mechanisms in the height-BC association.

Only energy restriction during the Economic Depression before and/or after the growth spurt had an effect on the mean adult-attained height, *i.e.* having an unemployed father during the Economic Depression resulted in a shorter stature compared to having an employed father. Exposure to energy restriction during the Hunger Winter or the War Years before and/or after the growth spurt was not associated with adult-attained height. We had foremost expected an effect of energy restriction on height for energy restriction during the Hunger Winter as this was the most extreme exposure. However exposure to the Hunger Winter, although severe, occurred relatively late in early life and was of relatively short duration, possibly enabling catch-up growth to take place.[39] The timing of exposure to early life energy restriction may be of importance in relation to it having a potential influence on adult-attained height, as the women were younger (0-23 years) during the Economic Depression as compared to the other exposures (12-28 years during the Hunger Winter and 7-28 years during the War Years).

With regard to energy restriction, previously the NLCS examined the association between early life energy restriction and breast cancer risk after 6.3 years of follow-up and did not observe any associations, neither with regard to the timing of the growth spurt. This may have to do with the lower number of cases. In addition, associations were not examined by hormone receptor-defined

subtypes.[25] In the current study, even though not all proxies of energy restriction influenced adult-attained height, all three proxies for energy restriction were associated with a (significantly) decreased risk of ER+ BC and PR+ BC in women who were exposed to energy restriction before and/or during the growth spurt, the group in which a potential effect on longitudinal growth was expected. The finding that both height and energy restriction, when occurring during a period in life in which adult-attained height is determined, are particularly associated with hormone receptor-positive BC risk supports the idea that common hormone-related growth-mechanisms may be involved. These findings also underline the notion that timing of exposure to early life energy restriction is an important factor to consider when studying BC risk.



The only study investigating energy restriction in relation to BC risk by hormone receptor subtypes reported an increased risk of ER-PR- BC, ER-PR+ BC, and ER+PR-BC, particularly in those exposed after birth (aged 0-3 years), while no association was reported for ER+PR+ BC. It should be noted, however, that the energy restriction (*i.e.*, China's Great famine) was quite extreme in this study as almost everyone experienced severe hunger during this famine and over 3% of the total population died as a result of the famine.[40,41] For early life energy restriction during the Hunger Winter, during which many women living in a western city also experienced severe hunger, we also observed an increased ER- BC risk, which seemed independent of timing of exposure to energy restriction with regard to the growth spurt. We speculate on the basis of these findings that this increased risk may in part have to do with the nature of the exposure, as both findings relate to severe energy restriction. Based on animal experimental models, we expected to find a decreased postmenopausal BC risk[42] that might be dose-dependent, though evidence also exists for a transition phase of the energy restriction effect: there may be a reversal of the effect from an increased to a decreased life- and health span at some level of energy restriction.[43,44] Energy intake reduction up to 65% improves the life- and health span in rodents, most noticeably by reducing the incidence of multiple forms of cancer. Yet, it has been suggested that an energy intake reduction of more than 65% may not impose the same health benefits regarding longevity.[13] With regard to early life energy restriction in animal models, the number of studies on energy intake reduction of more than 65% has been limited. The effects of extreme versus more moderate energy restriction on cancer risk are not clear yet but it is possible that the effects may differ with respect to the risk of (subtypes of) postmenopausal BC. However, residual confounding such as stress,[45] malnutrition and comorbidities related to

these severe famines[46-48] may also be partly responsible for the observed positive associations.

Strengths of the present study include the population-based prospective design and long follow-up, yielding large case numbers and making selection and information bias unlikely. Importantly, the elaborate available baseline information enabled us to adjust for a large set of relevant confounders, such as a number of reproductive factors, which are relevant for studying associations of adult-attained height and early life energy restriction with postmenopausal BC risk. In addition, information on ER+/- and PR+/- BC status in the NLCS was available for a relatively high percentage of BC cases (59% and 47%, respectively) and given the relatively large number of cases, we were able to conduct analyses on separate as well as combined ER+/- and PR+/-BC endpoints. Although the number of cases in some subgroup analyses were still small, this is the largest study to-date. It seems that taking into account heterogeneity between hormone receptor subtypes is of importance given that the observed associations with both adult-attained height and early life energy restriction differed for hormone receptor-positive and hormone receptor-negative BC subtypes. In the meta-analysis by our group, we were only able to investigate the risk of BC overall and observed a 28% increased BC risk for severe transient early life energy restriction.[13] Since the distribution of hormone receptor subtypes among BC cases may differ between study populations, and because associations with energy restriction seem to differ for the different BC subtypes, this may have affected the strength and direction of the observed association of early life energy restriction with the risk of BC overall in our meta-analysis. With regard to early life energy restriction, it should be mentioned that this is a unique exposure available within only a few cohorts worldwide.[13] Proxy measures were used to estimate energy restriction since information on individual food intake for the NLCS cohort was not available for the three periods of energy restriction, which may have resulted in some exposure misclassification. Nevertheless, any misclassification is likely to be non-differential, as individuals were still at risk for cancer at baseline when reporting on energy restriction via the proxy measures, which makes attenuation of hazard ratios most likely.

In conclusion, adult-attained height and early life energy restriction before and/or during the growth spurt were both associated with hormone receptor-positive BC risk, in the direction as expected, indicating critical exposure windows for hormonal growth-related mechanisms related to BC.

# REFERENCES


1 Wiren, S. *et al.* Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes Control* **25**, 151-159, doi:10.1007/s10552-013-0317-7 (2014).

2 World Cancer Research Fund / American Institute for Cancer Research. Continuous Update Project Report. Food, Nutrition, Physical Activity, and the Prevention of Breast Cancer. 2010.

3 van den Brandt, P. A. *et al.* Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol* **152**, 514-527 (2000).

4 Schonbeck, Y. *et al.* The world's tallest nation has stopped growing taller: the height of Dutch children from 1955 to 2009. *Pediatr Res* **73**, 371-377, doi:10.1038/pr.2012.189 (2013).

5 Cousminer, D. L. *et al.* Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. *Hum Mol Genet* **22**, 2735-2747, doi:10.1093/hmg/ddt104 (2013).

6 Canchola, A. J. *et al.* Body size and the risk of postmenopausal breast cancer subtypes in the California Teachers Study cohort. *Cancer Causes Control*, doi:10.1007/s10552-012-9897-x (2012).

7 Ritte, R. *et al.* Height, age at menarche and risk of hormone receptor-positive and -negative breast cancer: a cohort study. *Int J Cancer* **132**, 2619-2629, doi:10.1002/ijc.27913 (2013).

8 Borgquist, S., Jirstrom, K., Anagnostaki, L., Manjer, J. & Landberg, G. Anthropometric factors in relation to different tumor biological subgroups of postmenopausal breast cancer. *Int J Cancer* **124**, 402-411, doi:10.1002/ijc.23850 (2009).

9 Fagherazzi, G., Vilier, A., Boutron-Ruault, M. C., Clavel-Chapelon, F. & Mesrine, S. Height, sitting height, and leg length in relation with breast cancer risk in the E3N cohort. *Cancer Epidemiol Biomarkers Prev* **21**, 1171-1175, doi:10.1158/1055-9965.EPI-12-0130 (2012).

10 Bertrand, K. A., Gerlovin, H., Bethea, T. N. & Palmer, J. R. Pubertal growth and adult height in relation to breast cancer risk in African American women. *Int J Cancer*, doi:10.1002/ijc.31019 (2017).

11 Ruder, E. H., Dorgan, J. F., Kranz, S., Kris-Etherton, P. M. & Hartman, T. J. Examining breast cancer growth and lifestyle risk factors: early life, childhood, and adolescence. *Clin Breast Cancer* **8**, 334-342, doi:10.3816/CBC.2008.n.038 (2008).

12 Okasha, M., McCarron, P., Gunnell, D. & Smith, G. D. Exposures in childhood, adolescence and early adulthood and breast cancer risk: a systematic review of the literature. *Breast Cancer Res Treat* **78**, 223-276 (2003).

13 Elands, R. J. *et al.* A Systematic Literature Review and Meta-Regression Analysis on Early-Life Energy Restriction and Cancer Risk in Humans. *Plos One* **11**, e0158003, doi:10.1371/journal.pone.0158003 (2016).

14 Elias, S. G. *et al.* Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. *Int J Cancer* **108**, 628-630, doi:10.1002/ijc.11584 (2004).

15 Dahlmann, N. & Petersen, K. Influences of environmental conditions during infancy on final body stature. *Pediatr Res* **11**, 695-700, doi:10.1203/00006450-197705000-00014 (1977).

16 Angell-Andersen, E. *et al.* The association between nutritional conditions during World War II and childhood anthropometric variables in the Nordic countries. *Ann Hum Biol* **31**, 342-355, doi:10.1080/03014460410001685304 (2004).

17 Brundtland, G. H., Liestol, K. & Walloe, L. Height, weight and menarcheal age of Oslo schoolchildren during the last 60 years. *Ann Hum Biol* **7**, 307-322 (1980).
18 Nilsen, T. I. & Vatten, L. J. Adult height and risk of breast cancer: a possible effect of early nutrition. *Br J Cancer* **85**, 959-961, doi:10.1038/sj.bjc.6691946 (2001).
19 Li, C. I., Littman, A. J. & White, E. Relationship between age maximum height is attained, age at menarche, and age at first full-term birth and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* **16**, 2144-2149, doi:10.1158/1055-9965.EPI-07-0242 (2007).
20 Ahlgren, M., Melbye, M., Wohlfahrt, J. & Sorensen, T. I. Growth patterns and the risk of breast cancer in women. *N Engl J Med* **351**, 1619-1626, doi:10.1056/NEJMoa040576 (2004).
21 Velie, E. M., Nechuta, S. & Osuch, J. R. Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. *Breast Dis* **24**, 17-35 (2005).
22 De Stavola, B. L. *et al.* Childhood growth and breast cancer. *Am J Epidemiol* **159**, 671-682 (2004).
23 Alimujiang, A. *et al.* The association between China's Great famine and risk of breast cancer according to hormone receptor status: a hospital-based study. *Breast Cancer Res Treat* **160**, 361-369, doi:10.1007/s10549-016-3994-6 (2016).
24 van den Brandt, P. A., Dirx, M. J., Ronckers, C. M., van den Hoogen, P. & Goldbohm, R. A. Height, weight weight change, and postmenopausal breast cancer risk: The Netherlands Cohort Study. *Cancer Causes Control* **8**, 39-47 (1997).
25 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Diet in adolescence and the risk of breast cancer: results of the Netherlands Cohort Study. *Cancer Causes Control* **10**, 189-199 (1999).
26 van den Brandt, P. A. *et al.* A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol* **43**, 285-295 (1990).
27 Goldbohm RA, v. d. B. P., Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. *Tijdschr Soc Gezondheidsz*, 80-84 (1994).
28 Casparie, M. *et al.* Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* **29**, 19-24 (2007).
29 Burger G., S. H., Drummond J. Malnutrition and starvation in Western Netherlands, September 1944 to July 1945. Part I and II. . *The Hague, the Netherlands: General State Printing Office* (1948).
30 Trienekens, G. Tussen ons volk en de honger. De voedelvoorziening, 1940-45. (English translation: Between our nation and the hunger. The food supply, 1940-45). *Utrecht, the Netherlands: Stichting Matrijs* (1985).
31 J. Breunis The Annals of the American Academy of Political and Social Science, Vol. 245, The Netherlands during German Occupation (May, 1946), pp. 87-92.
32 Paternoster, L. *et al.* Adult height variants affect birth length and growth rate in children. *Hum Mol Genet* **20**, 4069-4075, doi:10.1093/hmg/ddr309 (2011).
33 Schouten, L. J., van Dijk, B. A., Lumey, L. H., Goldbohm, R. A. & van den Brandt, P. A. Energy restriction during childhood and early adulthood and ovarian cancer risk. *Plos One* **6**, e27960, doi:10.1371/journal.pone.0027960 (2011).

34 Barlow, W. E. Robust variance estimation for the case-cohort design. *Biometrics* **50**, 1064-1072 (1994).
35 Schoenfeld, D. Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239-241 (1982).
36 Sperrin, M., Marshall, A. D., Higgins, V., Renehan, A. G. & Buchan, I. E. Body mass index relates weight to height differently in women and older adults: serial cross-sectional surveys in England (1992-2011). *J Public Health (Oxf)* **38**, 607-613, doi:10.1093/pubmed/fdv067 (2016).
37 Droyvold, W. B. *et al.* Change in height, weight and body mass index: Longitudinal data from the HUNT Study in Norway. *Int J Obes (Lond)* **30**, 935-939, doi:10.1038/sj.ijo.0803178 (2006).
38 Zhang, B. *et al.* Height and Breast Cancer Risk: Evidence From Prospective Studies and Mendelian Randomization. *J Natl Cancer Inst* **107**, doi:10.1093/jnci/djv219 (2015).
39 Widen, E. *et al.* Distinct variants at LIN28B influence growth in height from birth to adulthood. *Am J Hum Genet* **86**, 773-782, doi:10.1016/j.ajhg.2010.03.010 (2010).
40 Kane, P. The demography of famine. *Genus* **43**, 43-58 (1987).
41 Song, S. Mortality consequences of the 1959-1961 Great Leap Forward famine in China: Debilitation, selection, and mortality crossovers. *Soc Sci Med* **71**, 551-558, doi:10.1016/j.socscimed.2010.04.034 (2010).
42 Dirx, M. J., Zeegers, M. P., Dagnelie, P. C., van den Bogaard, T. & van den Brandt, P. A. Energy restriction and the risk of spontaneous mammary tumors in mice: a meta-analysis. *Int J Cancer* **106**, 766-770, doi:10.1002/ijc.11277 (2003).
43 Roberts, S. B. & Speakman, J. Update on human calorie restriction research. *Adv Nutr* **4**, 563-564, doi:10.3945/an.113.004317 (2013).
44 Speakman, J. R. & Mitchell, S. E. Caloric restriction. *Mol Aspects Med* **32**, 159-221, doi:10.1016/j.mam.2011.07.001 (2011).
45 Vin-Raviv, N., Dekel, R., Barchana, M., Linn, S. & Keinan-Boker, L. World War II-related post-traumatic stress disorder and breast cancer risk among Israeli women: a case-control study. *Int Psychogeriatr* **26**, 499-508, doi:10.1017/S1041610213002081 (2014).
46 Bercovich, E., Keinan-Boker, L. & Shasha, S. M. Long-term health effects in adults born during the Holocaust. *Isr Med Assoc J* **16**, 203-207 (2014).
47 Koupil, I. *et al.* Blood pressure, hypertension and mortality from circulatory disease in men and women who survived the siege of Leningrad. *Eur J Epidemiol* **22**, 223-234, doi:10.1007/s10654-007-9113-6 (2007).
48 Huang, C., Li, Z., Wang, M. & Martore, R. Early Life Exposure to the 1959-1961 Chinese Famine Has Long-Term Health Consequences. *Journal of Nutrition* **140**, 1874-1878, doi:10.3945/jn.110.121293 (2010).

# Supplementary data

**Supplemental table 1. Hazard ratios and 95% confidence intervals for the association between adult-attained height and postmenopausal BC risk by combined hormone receptor subtypes in the Netherlands Cohort Study, 1986-2006**

| | | ER+PR+ | | | ER+PR- | | | ER-PR+ | | | ER-PR- | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PY | N | HR[a] | 95%CI | N | HR[a] | 95%CI | N | HR[a] | 95%CI | N | HR[a] | 95%CI |
| Height (cm) | | | | | | | | | | | | | |
| Tertile 1 (range: 132-163cm)[b] | 11,424 | 244 | 1 | ref | 67 | 1 | ref | 5 | 1 | ref | 80 | 1 | ref |
| Tertile 2 (range: 164-168cm)[b] | 10,609 | 259 | 1.15 | (0.92-1.43) | 75 | 1.19 | (0.84-1.71) | 7 | 1.53 | (0.48-4.90) | 54 | 0.72 | (0.49-1.05) |
| Tertile 3 (range: 169-198cm)[b] | 8,595 | 215 | 1.17 | (0.93-1.48) | 77 | 1.56 | (1.08-2.26) | 8 | 2.21 | (0.62-7.95) | 49 | 0.77 | (0.51-1.16) |
| p for trend | | | 0.18 | | | 0.02 | | | 0.22 | | | 0.18 | |
| Continuous per 5 cm | 30,629 | 718 | 1.07 | (0.99-1.16) | 219 | 1.12 | (0.99-1.28) | 20 | 1.27 | (0.81-2.01) | 183 | 0.96 | (0.84-1.10) |

Abbreviations: CI, confidence interval; ER+, estrogen receptor positive; ER-, estrogen receptor negative; HR, hazard ratio; PR+, progesterone receptor positive; PR-, progesterone receptor negative; PY, person-years.

[a] Adjusted for age (y), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency and smoking duration (number of cigarettes per day and number of years, respectively; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menarche (y), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever),and postmenopausal hormone-replacement therapy (never, ever).

[b] Tertiles were based on the distribution in the subcohort.

**Supplemental table 2. Hazard ratios and 95% confidence intervals for the association between exposure to early life energy restriction and postmenopausal BC risk by combined hormone receptor subtypes, including stratification by timing of the growth spurt[a] in the Netherlands Cohort Study, 1986-2006**

| | Postmenopausal breast cancer risk | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All women | | | | Women with energy restriction before/during the growth spurt[a] | | | | Women with energy restriction after the growth spurt[a] | | | |
| **Proxies for energy restriction by cancer endpoint** | **PY** | **N** | **HR[b]** | **95%CI** | **PY** | **N** | **HR[b]** | **95%CI** | **PY** | **N** | **HR[b]** | **95%CI** |
| *ER+PR+* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 420 | 1 | (ref) | 3,042 | 73 | 1 | (ref) | 13,506 | 347 | 1 | (ref) |
| Western rural | 4,150 | 93 | 0.88 | (0.66-1.17) | 791 | 17 | 1.09 | (0.51-2.31) | 3,359 | 76 | 0.84 | (0.61-1.16) |
| Western city | 8,250 | 162 | 0.74 | (0.59-0.93) | 1,603 | 16 | 0.23 | (0.10-0.54) | 6,647 | 146 | 0.83 | (0.65-1.07) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 251 | 1 | (ref) | 3,899 | 98 | 1 | (ref) | 6,641 | 153 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 283 | 0.93 | (0.73-1.18) | 4,705 | 86 | 0.60 | (0.39-0.91) | 7,681 | 197 | 1.15 | (0.84-1.57) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 635 | 1 | (ref) | 25,078 | 596 | 1 | (ref) | 1,169 | 39 | 1 | (ref) |
| Unemployed | 3,151 | 54 | 0.69 | (0.49-0.96) | 2,900 | 52 | 0.74 | (0.53-1.04) | 251 | 2 | —[c] | —[c] |
| *ER+PR-* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 127 | 1 | (ref) | 3,042 | 23 | 1 | (ref) | 13,506 | 104 | 1 | (ref) |
| Western rural | 4,150 | 37 | 1.26 | (0.84-1.88) | 791 | 10 | 2.53 | (0.88-7.28) | 3,359 | 27 | 1.12 | (0.71-1.79) |

The table Continues on the next page

**Table S2. continued**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Western city | 8,250 | 53 | 0.85 | (0.59-1.24) | 1,603 | 11 | 0.77 | (0.30-1.96) | 6,647 | 42 | 0.85 | (0.56-1.29) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 76 | 1 | (ref) | 3,899 | 30 | 1 | (ref) | 6,641 | 46 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 89 | 0.96 | (0.65-1.40) | 4,705 | 28 | 0.71 | (0.37-1.35) | 7,681 | 61 | 1.21 | (0.75-1.96) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 194 | 1 | (ref) | 25,078 | 187 | 1 | (ref) | 1,169 | 7 | 1 | (ref) |
| Unemployed | 3,151 | 17 | 0.69 | (0.40-1.20) | 2,900 | 17 | 0.76 | (0.44-1.32) | 251 | 0 | —[c] | —[c] |
| *ER-PR+* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 11 | 1 | (ref) | | | 1 | (ref) | | | 1 | (ref) |
| Western rural | 4,150 | 3 | —[c] | —[c] | 791 | 0 | —[c] | —[c] | 3,359 | 3 | —[c] | —[c] |
| Western city | 8,250 | 5 | 0.83 | (0.29-2.42) | 1,603 | 1 | —[c] | —[c] | 6,647 | 4 | —[c] | —[c] |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 6 | 1 | (ref) | 3,899 | 4 | 1 | (ref) | 6,641 | 2 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 8 | 1.16 | (0.30-4.41) | 4,705 | 2 | —[c] | —[c] | 7,681 | 6 | 2.78 | (0.28-27.71) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 17 | 1 | (ref) | 25,078 | 16 | 1 | (ref) | 1,169 | 1 | 1 | (ref) |
| Unemployed | 3,151 | 3 | —[c] | —[c] | 2,900 | 3 | —[c] | —[c] | 251 | 0 | —[c] | —[c] |

The table Continues on the next page

**Table S2. continued**

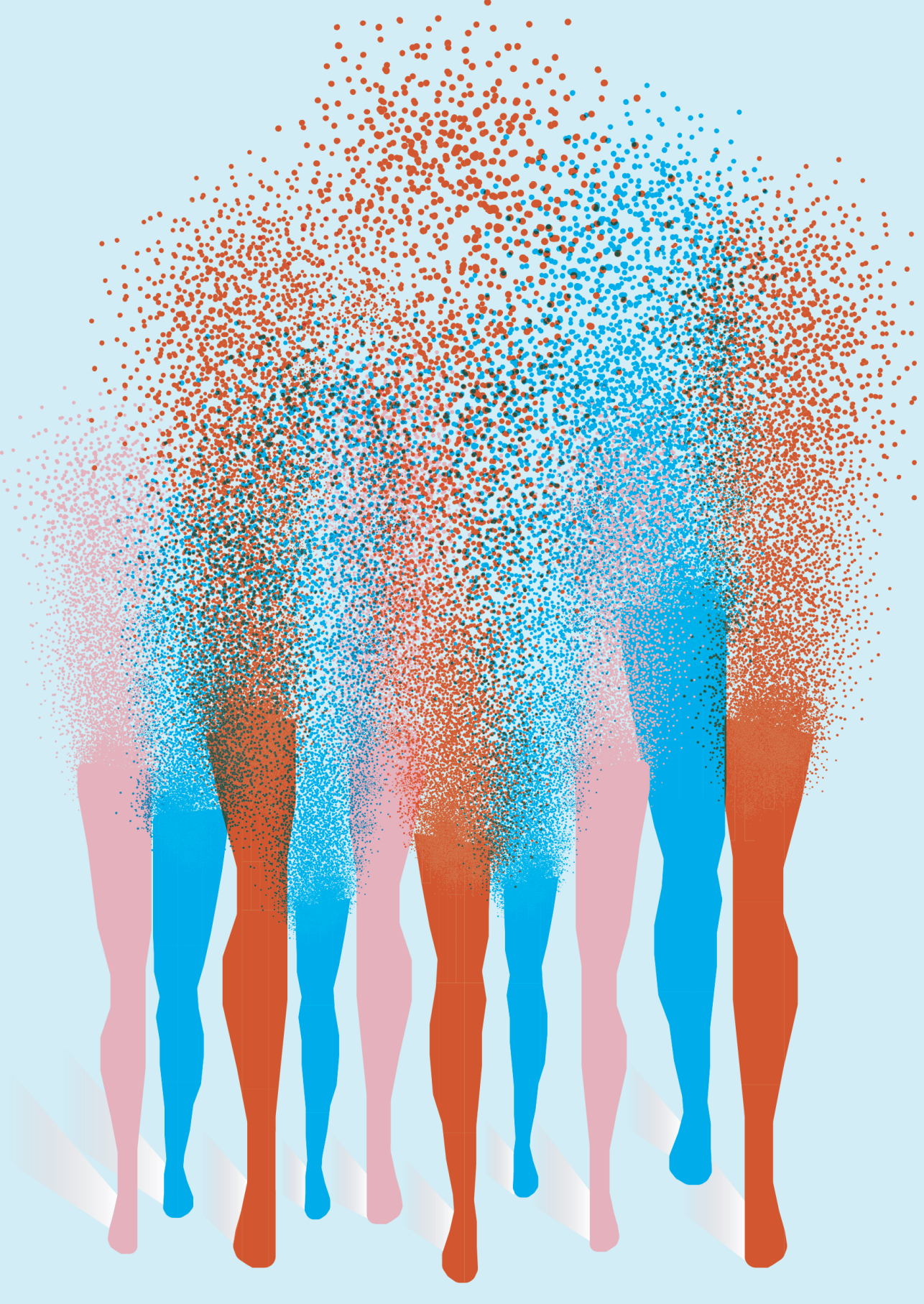
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *ER-PR-* | | | | | | | | | | | | |
| Residence during Hunger Winter (1944-45) | | | | | | | | | | | | |
| Non-western | 16,549 | 85 | 1 | (ref) | 3,042 | 17 | 1 | (ref) | 13,506 | 68 | 1 | (ref) |
| Western rural | 4,150 | 24 | 1.19 | (0.72-1.96) | 791 | 2 | –[c] | –[c] | 3,359 | 22 | 1.25 | (0.73-2.13) |
| Western city | 8,250 | 53 | 1.26 | (0.86-1.84) | 1,603 | 11 | 1.34 | (0.35-5.18) | 6,647 | 42 | 1.26 | (0.82-1.92) |
| Residence during War Years (1940-44) | | | | | | | | | | | | |
| Rural area in 1942 | 10,540 | 49 | 1 | (ref) | 3,899 | 16 | 1 | (ref) | 6,641 | 33 | 1 | (ref) |
| Urban area in 1942 | 12,387 | 78 | 1.25 | (0.82-1.89) | 4,705 | 30 | 1.19 | (0.56-2.49) | 7,681 | 48 | 1.28 | (0.75-2.19) |
| Job status father during Economic Depression (1932-40) | | | | | | | | | | | | |
| Employed | 26,247 | 157 | 1 | (ref) | 25,078 | 150 | 1 | (ref) | 1,169 | 7 | 1 | (ref) |
| Unemployed | 3,151 | 19 | 1.04 | (0.61-1.78) | 2,900 | 19 | 1.13 | (0.66-1.95) | 251 | 0 | –[c] | –[c] |

Abbreviations: CI, confidence interval; ER+, estrogen receptor positive; ER-, estrogen receptor negative; HR, hazard ratio; PR+, progesterone receptor positive; PR-, progesterone receptor negative; PY, person-years.

[a] The growth spurt was defined as: 2 years before reported age of menarche till 1 year after reported age of menarche.

[b] Adjusted for age (y), BMI (kg/m$^2$), energy intake (kcal/d), baseline non-occupational physical activity (≤30 min/d, >30-≤60 min/d, >60-90 min/d, >90 min/d), smoking status (never, former, current), smoking frequency and smoking duration (number of cigarettes per day and number of years, respectively; continuous, centered), alcohol intake (0, 0.1–29, ≥30 g/d), level of education (primary school, lower vocational school, intermediate vocational/high school, higher vocational school/ university), family history of BC (no, yes), history of benign breast disease (no, yes), age at menopause (y), age at first childbirth (nulliparous, >25y, ≤ 25y), parity (*n* children), oral contraceptive use (never, ever), and postmenopausal hormone-replacement therapy (never, ever).

[c] Estimate was unstable because of a low number of cases.



Chapter 4



# A Systematic SNP selection Approach to Identify Mechanisms Underlying Disease Aetiology: Linking Height to Postmenopausal Breast and Colorectal Cancer Risk

Rachel JJ Elands, Colinda CJM Simons, Mona Riemenschneider, Aaron Isaacs, Leo J Schouten, Bas A Verhage, Kristel Van Steen, Roger WL Godschalk, Piet A van den Brandt, Monika Stoll, Matty P Weijenberg - *Sci Rep. 2017 Jan 24;7:41034*

## Abstract

Data from GWAS suggest that SNPs associated with complex diseases or traits tend to co-segregate in regions of low recombination, harbouring functionally linked gene clusters. This phenomenon allows for selecting a limited number of SNPs from GWAS repositories for large-scale studies investigating shared mechanisms between diseases. For example, we were interested in shared mechanisms between adult-attained height and post-menopausal breast cancer (BC) and colorectal cancer (CRC) risk, because height is a risk factor for these cancers, though likely not a causal factor. Using SNPs from public GWAS repositories at *p*-values < $1x10^{-5}$ and a genomic sliding window of 1 mega base pair, we identified SNP clusters including at least one SNP associated with height and one SNP associated with either post-menopausal BC or CRC risk (or both). SNPs were annotated to genes using HapMap and GRAIL and analysed for significantly overrepresented pathways using ConsensuspathDB. Twelve clusters including 56 SNPs annotated to 26 genes were prioritised because these included at least one height- and one BC risk- or CRC risk-associated SNP annotated to the same gene. Annotated genes were involved in Indian hedgehog signalling (*p*-value = $7.78x10^{-7}$) and several cancer site-specific pathways. This systematic approach identified a limited number of clustered SNPs, which pinpoint potential shared mechanisms linking together the complex phenotypes height, post-menopausal BC and CRC.

## Introduction

Knowledge on single nucleotide polymorphisms (SNPs) and gene-environment interactions associated with complex diseases provides insights into underlying etiologic mechanisms [1, 2]. Genome-wide gene-environment interaction studies have typically been applying two-step approaches that are aimed at increasing power. Two-step genome-wide gene-environment interaction studies often utilise a SNP reduction step, in which the number of SNPs to include in the analysis is reduced [3]. The SNPs are subsequently tested for interaction, limiting multiple testing. However, for large-scale epidemiological studies with exhaustive biosamples from which DNA is not immediately suitable for genome-wide platforms, *e.g.* DNA from nails, the only option is platforms allowing genotyping of a limited number of SNPs. For example, we have previously genotyped toenail DNA using the Agena Bioscience™ MassARRAY® platform, which allows genotyping of a maximum of 40 SNPs at once in large-scale epidemiologic studies [4]. Therefore, an alternative systematic strategy is needed to reduce the number of relevant SNPs for studying disease aetiology through, for example, gene-environment interactions. Data from genome-wide association studies (GWAS) suggest that SNPs associated with complex diseases or traits are not randomly distributed across the genome but tend to co-segregate in regions of low recombination, harbouring functionally linked gene clusters [5]. Such an enrichment of loci associated with complex traits or diseases has been observed throughout the human genome [5] and offers an opportunity to SNP reduction.

Approaches for gene-environment interaction studies differ according to study objective. Searching for genetic causes of disease is nowadays generally an agnostic approach. In gene-environment-wide interaction studies, the starting point is also typically the genetic variation and how its interaction with the environment can contribute to the missing heritability [6]. Alternatively, studies aimed at understanding how the environment is associated with cancer risk are generally performed via a hypothesis-driven approach where the starting point is the environmental factor and the genetic variation is a time-independent biomarker of pathway involvement [2]. We were interested in the association between adult-attained height and cancer risk. Adult-attained height is an established risk factor for cancer risk at several sites; the most convincing evidence has been reported for post-menopausal breast cancer and colorectal cancer risk [7, 8]. For every 5 cm increase in height, post-menopausal breast cancer risk is reported to be increased by 7 to 11% [7, 9, 10] and colorectal cancer risk is increased by 6 to 11%

in women and 4 to 9% in men [8, 10, 11]. Adult-attained height in itself is probably not causally related to cancer, but rather a consistent marker for shared mechanisms determining both height and cancer risk, *e.g.* growth processes, which are influenced by factors such as growth promoting hormones and energy balance in early life [12]. Height is determined in the first 20 years of life by aggregated genetic [13] and environmental components [12], which determine linear growth but may also spur neoplastic growth later in life. Although adult-attained height may not be a target for intervention to reduce cancer risk, understanding how height is associated with cancer risk is essential to expand our knowledge concerning the pathways that lead to cancer development later in life. To study shared mechanisms between height and post-menopausal breast and colorectal cancer risk, we have applied a systematic SNP reduction strategy based on existing GWAS repositories and based on the fact that SNPs associated with complex diseases or traits tend to co-segregate in regions of low recombination. This knowledge was taken forward and we sought for clusters that included both height- and either postmenopausal breast cancer- or colorectal cancer-associated SNPs (or both) by comprehensively overlaying GWAS for these endpoints.

# Methods

## Search strategy

SNPs from the publically available manually curated National Human Genome Research Institute (NHGRI) Catalog of published GWAS [14] and the Johnson and O'Donnell database [15] associated with either height, post-menopausal breast or colorectal cancer risk were selected if these had a $p$-value < $1x10^{-5}$, a minor allele frequency (MAF) ≥ 1% in Caucasians, and were added to the catalogues up to June, 2014. Selected SNPs also included SNPs from meta-analyses on GWAS, which may have included SNPs with a $p$-value < $1x10^{-5}$ that did not reach this threshold in individual GWAS. The $p$-value cut-off for the selection of SNPs is a rather liberal value given the focus on genetic variation that tags mechanisms important for the multiple phenotypes of interest, in this case, height, post-menopausal breast cancer and colorectal cancer. Therefore, allowing a liberal $p$-value threshold permits one to identify clustered GWAS SNPs for a combination of different traits or diseases rather than clustered GWAS SNPs for a single phenotype. Genome-wide significant common variants ($p$-value < $5x10^{-8}$) and common variants that do not reach this criterion explain substantially large amounts of the heritability of complex traits and complex diseases; because SNPs below genome-wide

significance ($p$-value > $5x10^{-8}$) with marginal individual effect sizes may likely interact with other common SNPs and environmental components [16, 17]. SNPs identified in non-Caucasian populations were included if the corresponding MAF was ≥ 1% in Caucasians, for the reason that SNP-phenotype associations from different ancestries in independent GWAS might be informative to single out regions that link height to cancer risk. Including these SNPs from GWAS with other ancestries will also make our selection more comprehensive given that a number of SNPs may not yet have been explored in populations from Caucasian ancestry as a consequence of low signal resolution in older GWAS or because of differences in SNP coverage across genotyping platforms.

## Clustering methodology

Our clustering methodology was based on the assumption that GWAS SNPs associated with complex diseases or traits are not randomly distributed across the genome but tend to cluster in regions of low recombination [5]. Using a sliding window of 1 megabase pair (Mbp), genomic regions including at least one SNP from GWAS associated with height and one SNP from GWAS associated with either post-menopausal breast or colorectal cancer risk (or both) located within were designated as a SNP cluster. SNPs were clustered from the first height- or cancer risk-associated SNP that was identified from GWAS until no additional SNPs within the genomic sliding window of 1 Mbp could be found (**Figure 1**). Each cluster was assigned a unique cluster ID. The reason for implementing a relatively wide-ranging genomic sliding window (1 Mbp) was to allow for a sufficient number of SNPs, associated with multiple phenotypes, to cluster in regions of low recombination. We experimentally tested more conservative genomic sliding windows (0.1, 0.2, 0.3, and 0.5 Mbp), which resulted in identifying clusters with height- and breast cancer risk- or colorectal cancer risk-associated SNPs, but SNPs annotated to the same gene were not always in the same cluster anymore (which particularly affected large clusters with multiple SNPs annotated to the same gene). Furthermore, a few clusters were no longer identified. A wide-ranging genomic sliding window is preferable because the majority of GWAS SNPs reside in non-coding regions, potentially marking long-ranging disease-associated areas rather than pointing to individual genes. For example, 40.8% of SNPs from GWAS in DNAse I hypersensitive sites can be linked to target promotors over distances longer than 250 Kbp [18].

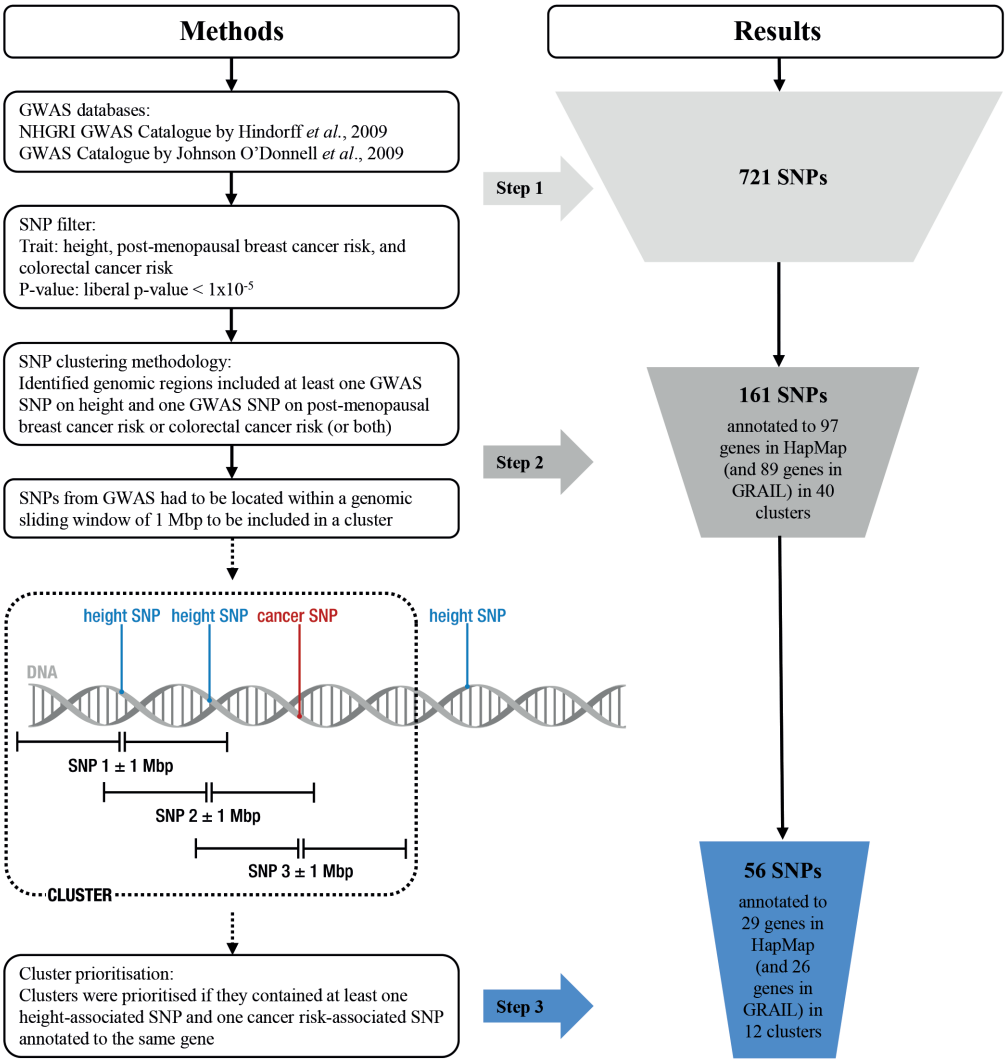



**Figure 1. Flow diagram with overview of SNP selection methodology and the corresponding results.**

SNPs from the clusters were geographically mapped to a gene according to HapMap release 37 and annotated to a gene according to "Gene Relationships Among Implicated Loci" (GRAIL) (https://www.broadinstitute.org/mpg/grail/). GRAIL accounts for the three-dimensional structure of the DNA, resulting in functional annotations. SNP clusters were prioritised when these contained at

least one height-associated SNP and one cancer risk-associated SNP that were mapped to the same gene according to the HapMap or GRAIL annotation (or both, allowing that HapMap and GRAIL may yield different annotations) or a combination of HapMap and GRAIL annotations. For each SNP in the prioritised set of clusters, the rs-number, mapped gene, publication information, SNP-phenotype information, the significance of the association, the effect size or beta-coefficient, confidence interval, ancestry and the risk allele (reported in the catalogues and from ENSEMBL) were collected. Within a cluster, pair-wise linkage disequilibrium (LD) was examined using SNAP version 2.2, (https://www.broadinstitute.org/mpg/snap/). Two or more SNPs in high pair-wise LD, *i.e.* $r^2 > 0.7$, marked redundant information within the cluster. Within LD pairs, SNPs with the lowest evidence for regulatory function annotation were excluded, but only if the cluster criteria were not violated. Ensembl Genome Browser was used to determine the genomic region of the SNPs and to identify whether these were localised in a regulatory region [19]. Regulatory functional annotation of SNPs was evaluated using a ranking ranging from 1-6 provided by RegulomeDB (http://www.regulomedb.org/) [20]. The ranking is based on the overlap of existing functional data including annotation to cis-expression quantitative trait loci (cis-eQTLs) and evidence for protein/transcription factor binding. SNPs that were likely linked to the expression of a gene target (cis-eQTLs) were assigned the highest possible ranking, i.e. scores 1a-1f, in RegulomeDB. SNPs that likely only affected protein binding were ranked lower (scores 2-3) and SNPs, for which there was minimal binding evidence (rank 4-6) or for which no evidence was available (score 0) were assigned the lowest evidence for regulatory function in RegulomeDB. The rationale to prioritise SNPs on the basis of regulatory information was derived from the knowledge that a significant number of SNPs associated with quantitative traits and common diseases in GWAS are concentrated in non-coding regulatory DNA sequences, therefore it is likely that regulatory processes underlie the relation between a SNP from GWAS and a phenotype [18, 21].

## Biological interpretation: gene set over-representation analyses

The gene annotations for the different SNPs in the resulting prioritised set of clusters, were imported to ConsensusPathDB (http://consensuspathdb.org/) [22] to conduct gene set over-representation analyses. In these analyses, pathways and gene ontology (GO) categories were tested for over-representation in the uploaded gene set. We primarily based these analyses on functional annotations from GRAIL. Tests were based on the hypergeometric test with a *p*-value cutoff

set to 0.01. Multiple testing was accounted for and the *q*-value threshold was set at 0.05. Pathway over-representation analyses and GO-over-representation analyses were performed for all clusters combined as well as separately for clusters including height- and post-menopausal breast cancer risk-associated SNPs and clusters including height- and colorectal cancer risk-associated SNPs.

## Results

An overview of the selection steps and the corresponding output is shown in **Figure 1**. The NHGRI Catalog included 1751 curated publications with 11,912 SNPs and the Johnson and O' Donnell database contained 56,411 SNPs from 118 articles. After selecting SNPs on the basis of the *p*-value ($p < 1x10^{-5}$) and MAF (≥ 1% in Caucasians) and filtering out duplicates, due to multiple associations in GWAS, we started clustering with 721 SNPs from both GWAS repositories. 514 SNPs were associated with height, 157 SNPs were associated with post-menopausal breast cancer risk and 50 SNPs were associated with colorectal cancer risk. None of the individual SNPs were associated with multiple phenotypes, *i.e.* height, post-menopausal breast cancer risk and/or colorectal cancer risk. Using the clustering method with a genomic sliding window of 1 Mbp, 40 clusters containing altogether 161 SNPs annotated to 97 genes on the basis of HapMap and 89 genes on the basis of GRAIL (9 SNPs could not be annotated) were formed, each including at least one SNP associated with height and one SNP associated either with post-menopausal breast or colorectal cancer risk (see Table S1). No SNP clusters were identified with combinations of SNPs that were associated with height, and both post-menopausal breast and colorectal cancer risk.

Twelve clusters containing altogether 56 SNPs, annotated to a total of 29 genes in HapMap and 26 genes in GRAIL (five SNPs could not be annotated), were prioritised as these clusters contained at least one height-associated SNP and one cancer risk-associated SNP that were annotated to the same gene. HapMap and GRAIL SNP-gene annotations were the same for 64.7% of the cases where both annotations were available (n = 51). Characteristics of the SNPs in the 12 prioritised SNP clusters are shown in Table 1 and S1. Eight SNPs in five of the prioritised clusters were eliminated from the total of 56 SNPs, leading to 48 SNPs in the prioritised clusters, due to the fact that these SNPs were in high LD ($r^2 > 0.7$) with another SNP in the same cluster, therefore these SNPs were likely to tag redundant

information. Of the 12 prioritised clusters, 8 clusters included 19 height- and 14 post-menopausal breast cancer risk-associated SNPs and four clusters included 10 height- and five colorectal cancer-risk associated SNPs. Of the 33 SNPs in height-breast cancer clusters, 26 SNPs were annotated to the same gene in sets of two or more height- and breast cancer risk-associated SNPs, leading to 9 gene annotations: *ID4, ZMIZ1, MCHR1* (in GRAIL)/*MKL1* (in HapMap)*, ESR1, RAD51B, TNS1, TNP1, TET2* and *FAM46A.* Of the 15 SNPs in height-colorectal cancer clusters, 8 SNPs were annotated to the same gene in pairs of height- and colorectal cancer-risk associated SNPs, leading to the following four gene annotations: *BMP2, PITX1, DCBLD1 and BARX1*. One prioritised cluster, cluster ID 22, contained two genes, *i.e. TNS1* and *TNP1*, to which height- and breast cancer risk-associated SNPs were annotated that were found associated in independent GWAS.

## Annotation of genomic region and regulatory function

According to Ensembl Genome Browser the majority of candidate SNPs ($n$ = 48) are located in introns ($n$ = 25) and in intergenic regions ($n$ = 17) (**Table 1**). The remaining SNPs were located in an enhancer ($n$ = 3), upstream of a gene ($n$ = 3), the promotor ($n$ = 3), an exon ($n$ = 3), or the promotor flanking region ($n$ = 1) (**Table 1**). According to RegulomeDB, 27 SNPs may affect transcription factor binding (score 1-5), of which five also affect the expression of a gene target, termed cis-eQTLs (score 1a-1f), and thus these had the highest regulatory evidence (**Table 1**).

**Table 1. Overview of the prioritised SNP clusters in which at least one height and one post-menopausal breast or colorectal cancer risk-associated SNP were annotated to the same gene as based on either HapMap or GRAIL, complemented by the SNP-annotation to biological regulatory function information and gene-annotation to enriched pathway and gene ontology categories.**

| Cluster ID | GWAS catalogue | | Genomic region based on Ensembl Genome Browser release 81 | Chromosome and cytogenicbond based on Ensembl Genome Browser release 81 | LD tag[c] | RegulomeDB | | | Mapped genein HapMap 37[g] | Annotated gene in GRAIL[h] | ConsensusPathDB analyses | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SNP ID[a] | Phenotype[b] | | | | Score[d] | Cis-eQTL[e] | Transcripition factor binding[f] | | | Gene ontology[i] | | | Pathway[j] |
| | | | | | | | | | | | #1 | #2 | #3 | |
| **Cluster ID 22** | rs13387042 | BC | intergenic | 2q | 0 | 0 | | | *TNP1* | *TNP1* | ✓ | ✓ | | |
| | rs2553026 | H | enhancer | | 0 | 0 | | | *TNP1* | NA | | | | |
| | rs1351164 | H | intron | | 0 | 0 | | | *TNS1* | NA | | | | |
| | rs16857609 | BC | intron | | 0 | 5 | | | *TNS1* | NA | | | | |
| | rs6435999 | H | intron | | 0 | 0 | | | *TNS1* | *DIRC3* | | | | |
| | rs3791950 | H | intergenic | | 0 | 2b | | *RUNX3, PAX5, TAF1* | *TNS1* | *TNS1* | | | | |
| | rs10187066 | H | intron | | 0 | 1f | *SLC11A1, CYP27A1* | | *ZNF142* | *STK36* | | | | Hedgehog signalling |
| | rs12470505*[a] | H | upstream gene | | 1 | 1f | *SLC23A3* | *USF1* | *CCDC108* | *IHH* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs1052483* | H | exon | | 1 | 1f | *SLC23A3* | | *NHEJ1* | *IHH* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs6724465* | H | intron | | 1 | 1f | *SLC23A3* | | *SLC23A3* | *IHH* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs16859517 | H | intergenic | | 0 | 5 | | | *SLC23A3* | *IHH* | ✓ | ✓ | ✓ | Hedgehog signalling |

The table Continues on the next page

**Table 1. Continued**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster ID 27** | rs9790517 | BC | intron | 4q | 0 | 0 | | | *TET2* | *TET2* | ✓ | ✓ |
| | rs10010325 | H | intron | | 0 | 0 | | | *TET2* | *TET2* | ✓ | ✓ |
| | rs6855629 | H | intron | | 0 | 6 | | | *TET2* | *EEF1AL7* | | |
| **Cluster ID 29** | rs526896* [a] | H | intergenic | 5q | 1 | 5 | | | *PITX1* | *PITX1* | ✓ | ✓ |
| | rs31198* | H | intron | | 1 | 5 | | | *PITX1* | *PITX1* | ✓ | ✓ |
| | rs647161 | CRC | intron | | 0 | 5 | | | *PITX1* | *PITX1* | ✓ | ✓ |
| **Cluster ID 32** | rs1047014 | H | upstream gene | 6p | 0 | 5 | | | *ID4* | *ID4* | ✓ | ✓ |
| | rs16882214 | BC | intergenic | | 0 | 0 | | | *ID4* | NA | | |
| **Cluster ID 33** | rs2322633 | H | intron | 6q | 0 | 6 | | | *BCKDHB* | *BCKDHB* | | |
| | rs310405 | H | intergenic | | 0 | 0 | | | *FAM46A* | *NA* | | |
| | rs17530068 | BC | intergenic | | 0 | 0 | | | *FAM46A* | *FAM46A* | | |
| **Cluster ID 34** | rs961764 | H | intergenic | 6q | 0 | 0 | | | *VGLL2* | *RFXDC1* | | |
| | rs2057314 | CRC | intron | | 0 | 4 | | *POLR2A, SPI1, TCF7L2, TCF12, NFIC, FOS* | *DCBLD1* | *DCBLD1* | | |
| | rs9285425 | H | intron | | 0 | 5 | | | *DCBLD1* | *DCBLD1* | | |
| **Cluster ID 7** | rs3757318* [a] | BC | intron | 6q | 1 | 2c | | *HNF4A, HNF4G* | *C6orf97* | *C6orf97* | | |
| | rs3734805* | BC | 3 prime UTR variant | | 1 | 0 | | | *C6orf97* | *C6orf97* | | |
| | rs2046210 | BC | intergenic | | 0 | 1f | *C6orf97* | | *ESR1* | *C6orf97* | | |
| | rs9383938 | BC | intron | | 0 | 5 | | *RFX3* | *ESR1* | *C6orf97* | | |
| | rs543650 | H | intron | | 0 | 0 | | | *ESR1* | *ESR1* | ✓ | ✓ ✓ |
| | rs9383951 | BC | intergenic | | 0 | 4 | | *GATA2, SETDB1* | *ESR1* | *ESR1* | ✓ | ✓ ✓ |
| | rs2982712 | H | intron | | 0 | 0 | | | *ESR1* | *ESR1* | ✓ | ✓ ✓ |

The table Continues on the next page

**Table 1. Continued**

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster ID 39** | rs10114408 | CRC | intergenic | 9q | 0 | 6 | | | *BARX1* | *BARX1* | ✓ | ✓ | | |
| | rs1257763 | H | intergenic | | 0 | 0 | | | *PTPDC1* | *BARX1* | ✓ | ✓ | | |
| | rs16910061 | H | upstream gene | | 0 | 5 | | *JUND* | *FBP2* | *FBP1* | ✓ | ✓ | | |
| | rs473902 | H | intron | | 0 | 3a | | *POL2RA* | *PTCH1* | *PTCH1* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs10512248 | H | promotor | | 0 | 6 | | | *PTCH1* | *PTCH1* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs2025151 | H | enhancer | | 0 | 1f | *HABP4* | *POL2RA, RUNX3* | *ZNF367* | *HABP4* | ✓ | ✓ | | |
| | rs10816533 | H | promotor | | 0 | 1f | *LOC642921* | | *ZNF510* | *ZNF782* | ✓ | ✓ | | |
| **Cluster ID 5** | rs704010 | BC | intron | 10q | 0 | 2b | | *POLR2A, SPI1, CTCF* | *ZMIZ1* | *ZMIZ1* | ✓ | ✓ | | |
| | rs7916441*[a] | H | enhancer | | 1 | 5 | | *GABPB1* | *ZMIZ1* | *ZMIZ1* | ✓ | ✓ | | |
| | rs780151* | H | intron | | 1 | 5 | | | *ZMIZ1* | *ZMIZ1* | ✓ | ✓ | | |
| | rs12355688 | BC | exon | | 0 | 4 | | *USF1, USF2* | *ZMIZ1* | *ZMIZ1* | ✓ | ✓ | | |
| | rs2145998* | H | intergenic | | 1 | 5 | | | *PPIF* | *ZMIZ1* | ✓ | ✓ | | |
| | rs941873 *[a] | H | promoter flanking | | 1 | 4 | | *HSF1, MAZ* | *ZCCHC24* | *ZMIZ1* | ✓ | ✓ | | |
| **Cluster ID 15** | rs2588809 | BC | intron | 14q | 0 | 0 | | | *RAD51B* | *RAD51B* | | | | |
| | rs1570106 | H | intron | | 0 | 0 | | | *RAD51B* | *RAD51B* | | | | |
| | rs999737 | BC | intron | | 0 | 6 | | | *RAD51B* | *RAD51B* | | | | |
| **Cluster ID 23** | rs961253 | CRC | intergenic | 20p | 0 | 5 | | | *FERMT1* | *FERMT1* | | | ✓ | |
| | rs967417* | H | intergenic | | 1 | 6 | | | *BMP2* | *BMP2* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs2145270* | H | intergenic | | 1 | 0 | | | *BMP2* | *BMP2* | ✓ | ✓ | ✓ | Hedgehog signalling |
| | rs2145272*[a] | H | intergenic | | 1 | 3a | | *STAT3, BCL11A, NFKB1, CHD2, EP300, IKZF1* | *BMP2* | *BMP2* | ✓ | ✓ | ✓ | Hedgehog signalling |

The table Continues on the next page

**Table 1. Continued**

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | rs4813802 | CRC | promotor | | 0 | 3a | *STAT3, CHD2, SETDB1, USF2, HNF4A, JUND, JUN, FOS, TRIM28, BACH1, TFAP2A, TFAP2C* | *BMP2* | *BMP2* | ✓ | ✓ | ✓ | Hedgehog signalling |
| **Cluster ID 25** | rs139909 | H | intron | 22q | 0 | 2b | *RUNX3, BATF, FOXM1,NFIC, ATF2, MTA3* | *TNRC6B* | *TNRC6B* | ✓ | ✓ | | |
| | rs5757949 | H | intron | | 0 | 5 | | *MKL1* | *MCHR1* | | | | |
| | rs6001930 | BC | intron | | 0 | 5 | *STAT3, CEBPB, EP300* | *MKL1* | *MCHR1* | | | | |

Abbreviations: eQTL; expression quantitative trait loci; GWAS, genome-wide association study; LD, linkage disequilibrium NA, data not available in GWAS catalogue; SNP, single nucleotide polymorphism.

[a] SNPs with the highest level of regulatory evidence were prioritised, indicated by the footnote ([a]). In cases were the regulatory evidence was equal, SNPs in high LD were prioritised according to the most significant *p*-value.

[b] Phenotype specifies whether a SNP derived from the GWAS catalogues by Hindorff et al., 2009 and Johnson O'Donnel et al., 2009 is associated with height (H), breast cancer risk (BC) or colorectal cancer risk (CRC).

[c] An LD tag equal to one denotes two or more SNPs within the same cluster that are in high LD ($r^2 > 0.7$).

[d] RegulomeDB score for the putative regulatory function of a SNP.

[e] Genes for which the SNP is a cis-eQTL according to RegulomeDB. (Cis-eQTLs are SNPs that are associated with mRNA expression of (a) nearby located gene(s).).

[f] Known transcription factor proteins that are binding to the genomic coordinates of the SNP according to RegulomeDB.

[g] SNPs were annotated to a gene using the physical mapping of a SNP to a gene according to HapMap.

[h] Gene annotations using GRAIL (http://software.broadinstitute.org/mpg/grail/) were based on gene relationships among the complete set of SNPs listed in this table (S1 table). In GRAIL, SNPs are annotated to genes by integrating the geographical location of a SNP derived from HapMap release 22with the biological data of a SNP obtained through text-mining using Pubmed 2014. GRAIL was set to correct for biases introduced by variable gene size when annotating the SNPs to genes. Large genes are more likely to have significant SNPs, and thus have a higher probability to be included in the regions that are being tested (Book: Computational Methods for Genetics of Complex Traits).

[i] Indicated with check-marks is whether the GRAIL gene annotation for a particular SNP contributed to the finding that the top three gene ontology terms, *i.e.* (#1) regulation of biosynthetic process (GO:009889), (#2) regulation of macromolecule metabolic process (GO:0060255), and (#3) epithelial cell proliferation (GO:0050673), were overrepresented in the total set of gene annotations from GRAIL (overrepresentation analyses were performed using ConsensusPathDB) .

[j] Indicates whether a gene mapped to a SNP is annotated to the overrepresented Indian hedgehog signalling pathway according to ConsensusPathDB.

## Pathway over-representation analyses

Pathway over-representation analysis based on the 26 gene annotations from GRAIL indicated the Indian hedgehog (Ihh) signalling pathway as the most significant overrepresented pathway (*p*-value = $7.78 \times 10^{-7}$) (based on the following genes: *BMP2, STK36, IHH, PTCH1*) (**Table 2**). Pathways that followed were ligand-receptor interactions (*IHH, PTCH1*) (*p*-value = $5.76 \times 10^{-5}$) and signalling in basal cell carcinoma (*BMP2*, *STK36*, *PTCH1*) (*p*-value = $6.73 \times 10^{-5}$) (**Table 2**). For comparison, when using the 29 HapMap gene annotations, the most significant overrepresented pathways were the Ihh signalling pathway, signalling in basal cell carcinoma, and the TGF-beta signalling pathway (data not shown).

A separate pathway over-representation analysis for genes annotated to SNPs that were associated with height or post-menopausal breast cancer risk also retrieved the Ihh pathway as the most overrepresented pathway (*STK36, IHH*) (*p*-value = $1.13 \times 10^{-4}$), as well as some distinct pathways, such as the ERBB4 signalling pathway (*ESR1, TNRC6B*) (*p*-value = $9.10 \times 10^{-3}$) and androgen receptor pathway (*ESR1, ZMIZ1*) (*p*-value = $9.02 \times 10^{-3}$) (**Table 2**). A separate pathway over-representation analysis for genes from clusters that contained SNPs associated with height or colorectal cancer risk, , indicated that the Ihh signalling pathway (*BMP2, PTCH1*) (*p*-value = $2.81 \times 10^{-4}$) and signalling in basal cell carcinoma (*BMP2, PTCH1*) (*p*-value = $2.53 \times 10^{-4}$) (**Table 2**) were overrepresented.

## Gene ontology over-representation analyses

A gene ontology term over-representation analysis, based on the 26 gene annotations from GRAIL, indicated the following top three most significantly overrepresented gene ontology terms for molecular and biological processes: regulation of biosynthetic process (*p*-value = $4.85 \times 10^{-6}$), regulation of macromolecule metabolic process (*p*-value = $2.85 \times 10^{-5}$) and epithelial cell proliferation (*p*-value = $3.29 \times 10^{-5}$) (**Table 3**).

**Table 2. Overrepresented pathways in prioritised SNP selection[a]**

| *Pathways* | *Set size* | *Number of genes from set in annotated gene list* | *Genes* | *p-value* | *q-value[b]* | *Pathway source* |
|---|---|---|---|---|---|---|
| **Overrepresented pathways using the genes annotated to the prioritised set of SNPs associated with height, post-menopausal breast and colorectal cancer risk.** | | | | | | |
| Hedgehog signalling pathway | 52 | 4 | *BMP2, STK36, IHH, PTCH1* | $7.78x10^{-7}$ | $2.08x10^{-5}$ | KEGG |
| Hedgehog signalling pathway | 16 | 3 | *STK36, IHH, PTCH1* | $1.49x10^{-6}$ | $2.08x10^{-5}$ | Wikipathways |
| Hedgehog | 25 | 3 | *STK36, IHH, PTCH1* | $6.06x10^{-6}$ | $5.56x10^{-5}$ | NetPath |
| Ligand-receptor interactions | 8 | 2 | *IHH, PTCH1* | $5.76x10^{-5}$ | $3.77x10^{-4}$ | Reactome |
| Basal cell carcinoma | 55 | 3 | *BMP2, STK36, PTCH1* | $6.73x10^{-5}$ | $1.63x10^{-3}$ | KEGG |
| HH-Core | 19 | 2 | *IHH, PTCH1* | $3.48x10^{-4}$ | $1.63x10^{-3}$ | Signalink |
| Signalling events mediated by the Hedgehog family | 23 | 2 | *IHH, PTCH1* | $5.14x10^{-4}$ | $2.06x10^{-3}$ | PID |
| Hedgehog, on, state | 42 | 2 | *IHH, PTCH1* | $1.41x10^{-3}$ | $4.93x10^{-3}$ | Reactome |
| Hedgehog signalling events mediated by Gli proteins | 50 | 2 | *STK36, PTCH1* | $2.24x10^{-3}$ | $6.97x10^{-3}$ | PID |
| Endochondral ossification | 64 | 2 | *IHH, PTCH1* | $3.83x10^{-3}$ | $1.07x10^{-3}$ | Wikipathways |
| TGF-beta signalling pathway | 80 | 2 | *BMP2, ID4* | $5.96x10^{-3}$ | $1.48x10^{-3}$ | KEGG |
| Signalling by Hedgehog | 87 | 2 | *IHH, PTCH1* | $6.41x10^{-3}$ | $1.48x10^{-3}$ | Reactome |
| Class B/2 (Secretin family receptors) | 88 | 2 | *IHH, PTCH1* | $6.87x10^{-3}$ | $1.48x10^{-3}$ | Reactome |
| **Overrepresented pathways using the genes annotated to the prioritised SNPs associated with height and post-menopausal breast cancer risk.** | | | | | | |
| Hedgehog signalling pathway | 16 | 2 | *STK36, IHH* | $1.13x10^{-4}$ | $1.35x10^{-3}$ | Wikipathways |
| Hedgehog | 25 | 2 | *STK36, IHH* | $2.81x10^{-4}$ | $1.68x10^{-3}$ | NetPath |
| Hedgehog signalling pathway | 52 | 2 | *STK36, IHH* | $1.18x10^{-3}$ | $4.70x10^{-3}$ | KEGG |
| Signalling by *ERBB4* | 153 | 2 | *ESR1, TNRC6B* | $9.02x10^{-3}$ | $2.19x10^{-2}$ | Reactome |
| Androgen receptor | 149 | 2 | *ESR1, ZMIZ1* | $9.14x10^{-3}$ | $2.19x10^{-2}$ | NetPath |
| **Overrepresented pathways using the genes annotated to the prioritised SNPs associated with height and colorectal cancer risk.** | | | | | | |
| Hedgehog signalling pathway | 52 | 2 | *BMP2, PTCH1* | $2.81x10^{-4}$ | $6.34x10^{-4}$ | KEGG |
| Basal cell carcinoma | 55 | 2 | *BMP2, PTCH1* | $2.53x10^{-4}$ | $6.34x10^{-4}$ | KEGG |

Abbreviations: SNP, single nucleotide polymorphism.

[a] Overrepresented pathways were retrieved using the SNP-gene annotations from GRAIL.
[b] The *p*-values are corrected for multiple testing using the false discovery rate method and are shown as *q*-values.

**Table 3. Top ten most significantly overrepresented gene-ontology terms in prioritised SNP selection[a]**

| GO terms[a] | Set size | Number of genes from set in annotated gene list | p-value | q-value[b] | Sub-analysis: height and breast cancer risk[c] | Sub-analysis: height and colorectal cancer risk[c] |
|---|---|---|---|---|---|---|
| GO:0009889 regulation of biosynthetic process | 4061 | 15 | $4.85 \times 10^{-6}$ | $6.21 \times 10^{-4}$ | ✓ | |
| GO:0060255 regulation of macromolecule metabolic process | 5358 | 16 | $2.85 \times 10^{-5}$ | $1.80 \times 10^{-3}$ | ✓ | |
| GO:0050673 epithelial cell proliferation | 323 | 5 | $3.29 \times 10^{-5}$ | $3.30 \times 10^{-2}$ | ✓ | ✓ |
| GO:0048754 branching morphogenesis of an epithelial tube | 170 | 4 | $4.55 \times 10^{-5}$ | $1.80 \times 10^{-3}$ | ✓ | ✓ |
| GO:0090304 nucleic acid metabolic process | 4893 | 15 | $5.61 \times 10^{-5}$ | $1.80 \times 10^{-3}$ | ✓ | |
| GO:0016070 RNA metabolic process | 4339 | 14 | $7.48 \times 10^{-5}$ | $1.81 \times 10^{-3}$ | ✓ | |
| GO:0061138 morphogenesis of a branching epithelium | 202 | 4 | $8.47 \times 10^{-5}$ | $1.81 \times 10^{-3}$ | ✓ | |
| GO:0048732 gland development | 407 | 5 | $9.38 \times 10^{-5}$ | $3.30 \times 10^{-3}$ | ✓ | ✓ |
| GO:0060322 head development | 678 | 6 | $10.40 \times 10^{-4}$ | $3.30 \times 10^{-3}$ | ✓ | |
| GO:0001763 morphogenesis of a branching structure | 213 | 4 | $10.50 \times 10^{-4}$ | $3.30 \times 10^{-3}$ | ✓ | |

Abbreviations GO, gene ontology; SNP, single nucleotide polymorphism.

[a] Overrepresentation analysis for GO terms were performed using using the SNP-gene annotations from GRAIL.

[b] The $p$-values are corrected for multiple testing using the false discovery rate method and are available as $q$-values.

[c] The check-mark indicates which of the top 10 GO-terms from the main GO overrepresentation analysis were also present in separate analyses for breast and colorectal cancer risk.

## Discussion

We present a systematic approach for epidemiologic studies to prioritise SNPs associated with multiple complex diseases or traits using all GWAS repository data publically available to elucidate aetiologic pathways. The clustering methodology in this approach relies on the assumption that SNPs from GWAS found associated with complex diseases or traits are not randomly distributed across the genome, but tend to cluster in regions of low recombination [5]. This allows for a systematic narrowing down of the genomic search field and we were able to identify clusters that were of relevance to the height-cancer association. Twelve clusters were identified that contained at least one height- and one cancer risk-associated SNP annotated to the same gene. Height- and post-menopausal breast cancer risk-associated SNPs (*n* = 33) clustered together in 8 clusters. In these, 26 SNPs were annotated to the same gene in sets of two or more height- and breast cancer risk-associated SNPs, leading to the following 9 gene annotations: *ID4, ZMIZ1, MCHR1* (in GRAIL)*/MKL1* (in HapMap)*, ESR1, RAD51B, TNS1, TNP1, TET2* and *FAM46A*. Height- and colorectal cancer risk-associated SNPs (*n* = 15) clustered together in four clusters. In these, 8 SNPs were annotated to the same gene in pairs of height- and colorectal cancer risk-associated SNPs, leading to the following four gene annotations: *BMP2, PITX1, DCBLD1*, and *BARX1*.

The SNP selection strategy proposed here can typically be used to identify shared mechanisms between multiple traits or diseases, using gene-environment interactions for example. A number of two-step methods have been developed based on genome-wide data prioritising relevant SNPs within the own study population and subsequently testing these SNPs for interactions [3, 6]. These existing strategies prioritise SNPs related to exposure in cases and controls [23] or SNPs related to the outcome [24]. The cocktail-method is an approach which combines features of two-step methods, the case-only design, and empirical Bayes techniques [25]. Still, these strategies inherently lead to a higher probability of type I error, because SNPs are prioritised based on a genome-wide scan in the own study population without replication of the result. This can be avoided by selecting SNPs from publically available GWAS data, independent of the own study population, and using the clustering methodology to identify genomic regions of importance in relation to the phenotypes of interest. For most SNP clusters marking these regions, there is no particular expectation that the set of SNPs associated with the phenotypes of interest are themselves causal variants. Rather, the clusters mark regions in the human genome, which correlate

with one or more causal variants. Therefore, the GWAS SNPs found in a single region likely tag similar mechanisms or causal variants and, in a way, may act as replication of the same result. These SNPs can then be taken forward to test for gene-environment interactions. The SNPs in the clusters may collectively point to pathways explaining the link between height and cancer risk. Previously, Mendelian randomization has been employed to make causal inferences regarding the link between height and colorectal cancer risk utilising genetic variants as a proxy for height. For example, Thrift *et al.* (2015) suggested a causal association between height-increasing alleles and a higher colorectal cancer risk in women, but further investigation was warranted in men [26]. An additional advantage of the clustering approach is that it is also particularly suitable for the investigation of several SNPs at once, all within one cluster, *e.g.* through the use of a genetic risk score, thereby accounting for multiple SNP effects and reducing the multiple testing problem.

Our SNP selection approach may also have some limitations. For example, the size of the genomic sliding window affected the cluster size and the number of clusters identified. Also, the method is reliant on published GWAS data which are not freely available at $p$-values $\geq 1x10^{-5}$ in the NHGRI GWAS Catalog and $p$-values $> 1x10^{-3}$ in the Johnson and O'Donnell database. Furthermore, the number of SNPs from GWAS on height is relatively high compared to the number of SNPs from GWAS on breast and colorectal cancer risk; this might have to do with the fact that anthropometric data such as height is available in most studies. Nevertheless, the observation that a number of pathways of relevance to both height, post-menopausal breast cancer risk, and colorectal cancer risk were found overrepresented among the genes annotated to the SNPs in the clusters suggests that this approach can reveal biologically relevant information.

The notion that specific genes [27, 28] and genetic variants [26, 29, 30] may be relevant for explaining the height-cancer association has been suggested previously. Our systematic SNP selection strategy showed the Ihh signalling pathway to be overrepresented as based on variants that lie in/near *BMP2, IHH, PTCH1, and STK36,* when basing gene annotations on GRAIL. Cross-talks have been suggested between the Ihh signalling pathway and the Transforming Growth Factor-beta (TGF-β) signalling pathway, which was found in overrepresentation analyses using HapMap gene annotations. Both pathways are of relevance to processes in growth plate regulation and the length of bones [31, 32] as well as tumour development [33, 34]. Few hypothesis-based candidate-gene studies have been performed on SNPs in Ihh signalling pathway genes and breast or colorectal cancer risk. SNPs in TGF-β

signalling pathway genes have been associated with increased breast cancer risk [35]. Moreover, it has been found that a high number of at-risk variants in genes in the TGF-β signalling pathway increased the risk of colon and rectal cancer [36]. That cross-talks between Ihh and TGF-β signalling pathways are important in linking height to cancer, is likely when considering other complex diseases such as coronary artery disease (CAD). Consistent with an inverse association between height and CAD, a recent study showed that genetically determined height, as based on 180 height-associated SNPs from the Genetic Investigation of Anthropometric Traits (GIANT) consortium (which were not found in GWAS on CAD), was inversely associated with CAD, possibly via BMP/TGF-β signalling [37]. Furthermore, interestingly, the basal cell carcinoma pathway is also significantly overrepresented in our results, which supports the previously reported height-basal cell cancer association [38].

A number of SNPs were annotated to genes that fall in unanticipated pathways. Even though these pathways were not identified in our pathway overrepresentation analysis, these SNPs may provide new clues about the mechanisms that influence growth in relation to adult-attained height and breast and colorectal cancer risk. For example, of interest may be the melanin-concentrating hormone receptor (*MCHR1)* gene, to which both height- and breast cancer risk-associated SNPs were annotated. Several studies have supported a role for *MCHR1* in the regulation of food consumption behaviour, energy expenditure and body weight [39, 40]. Previously, a cross-sectional study found that polymorphisms in the *MCHR1* gene were associated with differences in body composition and interacted with energy-related lifestyle factors [41]. Body fatness is, next to adult-attained height, a convincing risk factor for post-menopausal breast cancer [7]. Therefore, nutrient-sensing processes might be a common mechanism linking height and other anthropometric factors to breast cancer risk.

Unexpectedly, no clusters were identified that contained SNPs that were associated with all three phenotypes, *i.e.* height, post-menopausal breast cancer risk, and colorectal cancer risk. This might be explained by the fact that the *p*-value cut-off (*p*-value = $1x10^{-5}$) used for GWAS SNPs, although liberal, was not sufficiently liberal to find clusters that represented all three phenotypes. Likely, at even more liberal *p*-values, there is a higher probability of finding a shared component to complex traits, such as height and the risk of cancer, which may be involving thousands of common alleles with rather small effects [42]. Our results suggest that, in addition to a shared component, there may also be different mechanisms through which height influences post-menopausal breast and

colorectal cancer risk. The mechanisms identified linking height to colorectal cancer risk overlapped with those found in overall pathway overrepresentation analyses in this study and these may operate primarily through Ihh signalling. The mechanisms linking height to post-menopausal breast cancer risk may go through Ihh signalling as well as ERBB4 signalling and androgen receptor signalling. Both ERBB4 signalling [43, 44] and androgen receptor signalling [45, 46] are involved in mammary gland development. Future studies can utilise the SNPs in height-post-menopausal breast and height-colorectal cancer clusters to conduct mediation analyses between SNPs and specific cancer endpoints with height as a mediating factor or to perform interaction analyses between SNPs and height with specific cancer endpoints.

Finally, it is only fair to mention that our method is likely to pick up some degree of pleiotropic effects in terms of SNP effects or gene effects, especially considering our prioritisation step in which we prioritised clusters with at least one height- and one cancer risk-associated SNP. In this report, however, we focused on the instrumental value of the clusters in terms of future gene-environment interaction analyses or mediation analyses aimed at elucidating disease aetiology, rather than on trying to pinpoint pleiotropic SNPs or genes. Nevertheless, it is good to realise that several other methods exist that are aimed at identifying potential pleiotropic effects [47-49]. These methods may, in part, confirm the results at hand, when applied to the same topic. However, due to differences in input and methodology, it is likely that also different signals will be picked up. It is beyond the scope of this paper to identify all existing methods and validate these against each other, but we encourage future efforts in relation to this issue. Such efforts preferably need to include the use of simulated data in order to be able to draw conclusions about the extent to which different signals are picked up by different methods and about the extent to which different methods can distinguish between true signals and noise.

## Conclusion

We report a novel SNP selection approach to systematically restrict the number of SNPs for genotyping in large-scale studies aimed at elucidating aetiologic pathways. Our approach is of particular interest for studies with exhaustive bio-samples, in which a genome-wide approach is not feasible, and will reduce the costs of genotyping and the chance of false-positive findings. The SNPs identified can be used to, for example, study gene-environment interactions or to conduct mediation analyses. The novelty of this method is the comprehensive integration of publically available GWAS repositories on the basis of which SNPs associated

with multiple linked complex traits and diseases can be identified as these are hypothesised to cluster in regions of low recombination. Such SNPs may serve as time-independent biomarkers of pathway involvement to mechanistically underpin established associations. Of interest in this paper was the association between adult-attained height and the risk of post-menopausal breast and colorectal cancer, for which the Ihh signalling pathway was found to be potentially important. This pathway was also found in separate analyses for height-post-menopausal breast cancer and height-colorectal cancer clusters, but there may also be different biological mechanisms through which height is associated with post-menopausal breast as compared to colorectal cancer risk.

# References


1. Hunter, DJ, Altshuler, D, Rader, DJ. From Darwin's finches to canaries in the coal mine--mining the genome for new biology. *N Engl J Med.* 358(26):2760-3; DOI: 10.1056/NEJMp0804318 (2008)
2. Le Marchand, L, Wilkens, LR. Design considerations for genomic association studies: importance of gene-environment interactions. *Cancer Epidemiol Biomarkers Prev.*17(2):263-7; DOI: 10.1158/1055-9965.EPI-07-0402 (2008)
3. Hutter, CM, et al. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank report. *Genet Epidemiol.* 37(7):643-57; DOI: 10.1002/gepi.21756 (2013)
4. Hogervorst, J, et al. DNA from nails for genetic analyses in large-scale epidemiologic studies. *Cancer Epidemiol Biomarkers Prev.* 23(12):2703-12; DOI: 10.1158/055-9965.EPI-14-0552 (2014)
5. Preuss C, Riemenschneider M, Wiedmann D, Stoll M. Evolutionary dynamics of co-segregating gene clusters associated with complex diseases. PLoS One. 2012;7(5):e36205. doi: 10.1371/journal.pone.0036205.
6. Thomas, D. Gene--environment-wide association studies: emerging approaches. *Nat Rev Genet.* 11(4):259-72; DOI: 10.1038/nrg2764 (2010)
7. World Cancer Research Fund / American Institute for Cancer Research. Continuous Update Project Report. Food, Nutrition, Physical Activity, and the Prevention of Breast Cancer. http://www.wcrf.org/int/research-we-fund/continuous-update-project-findings-reports/breast-cancer (2010).
8. World Cancer Research Fund / American Institute for Cancer Research. Continuous Update Project Report. Food, Nutrition, Physical Activity, and the Prevention of Colorectal Cancer. http://www.wcrf.org/int/research-we-fund/continuous-update-project-findings-reports/colorectal-bowel-cancer (2011).
9. van den Brandt, PA, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol.* 152(6):514-27; (2000)
10. Wiren, S, et al. Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes Control.* 25(2):151-9; DOI: 10.1007/s10552-013-0317-7 (2014)
11. Wei, EK, et al. Comparison of risk factors for colon and rectal cancer. *Int J Cancer.* 108(3):433-42; DOI: 10.1002/ijc.11540 (2004)
12. Okasha, M, Gunnell, D, Holly, J, Davey Smith, G. Childhood growth and adult cancer. *Best Pract Res Clin Endocrinol Metab.* 16(2):225-41; DOI: 10.1053/beem.2002.0204 (2002)
13. Silventoinen, K, et al. Genetic regulation of growth from birth to 18 years of age: the Swedish young male twins study. *Am J Hum Biol.* 20(3):292-8; DOI: 10.1002/ajhb.20717 (2008)
14. Hindorff, LA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A.* 106(23):9362-7; DOI: 10.1073/pnas.0903103106 (2009)
15. Johnson, AD, O'Donnell, CJ. An open access database of genome-wide association results. *BMC Med Genet.* 10:6; DOI : 10.1186/1471-2350-10-6 (2009)
16. Yang, J, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 42(7):565-9; DOI: 10.1038/ng.608 (2010)

17. Zhang, G, et al. Finding missing heritability in less significant Loci and allelic heterogeneity: genetic variation in human height. *PLoS One* 7(12):e51211; DOI: 10.1371/journal.pone.0051211 (2012)
18. Maurano, MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337(6099):1190-5; DOI: 10.1126/science.1222794 (2012)
19. Cunningham, F, et al. Ensembl 2015. *Nucleic Acids Res.* 43(Database issue):D662-9; DOI: 10.1093/nar/gku1010 (2015)
20. Boyle, AP, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 22(9):1790-7; DOI: 10.1101/gr.137323.112 (2012)
21. Schaub, MA, Boyle, AP, Kundaje, A, Batzoglou, S, Snyder, M. Linking disease associations with regulatory information in the human genome. *Genome Res.* 22(9):1748-59; DOI: 10.1101/gr.136127.111 (2012)
22. Kamburov, A, Stelzl, U, Lehrach, H, Herwig, R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res.* ;41(Database issue):D793-800; DOI: 10.1093/nar/gks1055 (2013)
23. Murcray, CE, Lewinger, JP, Gauderman, WJ. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol.* 169(2):219-26; DOI: 10.1093/aje/kwn353 (2009)
24. Kooperberg, C, Leblanc, M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. *Genet Epidemiol.* 32(3):255-63; DOI: 10.1002/gepi.20300 (2008)
25. Hsu, L, et al. Powerful cocktail methods for detecting genome-wide gene-environment interaction. *Genet Epidemiol.* 36(3):183-94; DOI: 10.1002/gepi.21610 (2012).
26. Thrift, AP, et al. Mendelian randomization study of height and risk of colorectal cancer. *Int J Epidemiol.* 44(2):662-72; DOI: 10.1093/ije/dyv082 (2015)
27. Tripaldi, R, Stuppia, L, Alberti, S. Human height genes and cancer. *Biochim Biophys Acta.* 1836(1):27-41; DOI: 10.1016/j.bbcan.2013.02.002 (2013)
28. Stevens, A, et al. Human growth is associated with distinct patterns of gene expression in evolutionarily conserved networks. *BMC Genomics.*14:547; DOI: 10.1186/1471-2164-14-547 (2013)
29. Kitahara, CM, et al. Association between adult height, genetic susceptibility and risk of glioma. *Int J Epidemiol.* 41(4):1075-85; DOI: 10.1093/ije/dys114 (2012)
30. Thrift, AP, et al. Risk of esophageal adenocarcinoma decreases with height, based on consortium analysis and confirmed by Mendelian randomization. *Clin Gastroenterol Hepatol.* 12(10):1667-76 e1; DOI: 10.1016/j.cgh.2014.01.039 (2014)
31. van der Eerden, BC, Karperien, M, Wit, JM. Systemic and local regulation of the growth plate. *Endocr Rev.* 24(6):782-801; DOI: 10.1210/er.2002-0033 (2003).
32. Lui, JC, et al. Synthesizing genome-wide association studies and expression microarray reveals novel genes that act in the human growth plate to modulate height. *Hum Mol Genet.* 21(23):5193-201 DOI: 10.1093/hmg/dds347 (2012)
33. Fuxe, J, Vincent, T, Garcia de Herreros, A. Transcriptional crosstalk between TGF-beta and stem cell pathways in tumor cell invasion: role of EMT promoting Smad complexes. *Cell Cycle* 9(12):2363-74; DOI: 10.4161/cc.9.12.12050 (2010)
34. Hameetman, L, et al. Peripheral chondrosarcoma progression is accompanied by decreased Indian Hedgehog signalling. *J Pathol.* 209(4):501-11; DOI: 10.1002/path.2008 (2006)

# Supplementary data

**Table S1. SNP clusters formed using SNPs from the GWAS catalogues by Hindorff *et al*., 2009 [1] and Johnson O'Donnell *et al*., 2009 [2]; SNPs were associated with height, post-menopausal breast or colorectal cancer risk at a p-value < $10^{-5}$ and clusters were formed using a 1 mega base pair window and included at least one height- and one cancer risk-associated SNP**

| GWAS catalogue | | | HapMap Gene[a] | GRAIL Gene[b] | Cluster ID | LD tag[c] | Genomic region based on the Ensemble Browser | RegulomeDB | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SNP ID | P value | Phenotype | | | | | | Score[d] | eQTL[e] | Protein binding[f] |
| rs11548323 | $2x10^{-06}$ | Breast cancer | *WASF2* | *FCN3* | 1 | 0 | 3 prime UTR | 4 | | *CTCF* |
| rs11809207 | $6x10^{-08}$ | Height | *CATSPER4* | *CATSPER4* | 1 | 0 | Intron | 0 | | - |
| rs7532866 | $3x10^{-08}$ | Height | *LIN28AP1* | *LIN28* | 1 | 0 | Intron | 6 | | - |
| rs1387389 | $4x10^{-06}$ | Breast cancer | *PBX1* | *PBX1* | 2 | 0 | Intron | 5 | | *hCPE-R* |
| rs6670655 | $3x10^{-06}$ | Height | *PBX1* | *PBX1* | 2 | 0 | Non coding transcript exon variant | 4 | | *HNF4A* |
| rs1046934 | $2x10^{-31}$ | Height | *TSEN15* | *C1orf19* | 3 | 1 | Missense variant | 6 | | - |
| rs10911251 | $9x10^{-08}$ | Colorectal cancer | *LAMC1* | *LAMC1* | 3 | 0 | Intron | 6 | | - |
| rs2274432 | $8x10^{-09}$ | Height | *TSEN15* | *C1orf19* | 3 | 1 | Missense variant | 4 | | *E2F1* |
| rs3814333 | $2x10^{-13}$ | Height | *COLGALT2* | *C1orf19* | 3 | 0 | Upstream gene variant | 5 | | *CEBPB, CDX2* |
| rs756199 | $7x10^{-06}$ | Height | *COLGALT2* | *C1orf19* | 3 | 0 | Intron | 5 | | - |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| rs10508468 | 7x10$^{-05}$ | Breast cancer | *FRMD4A* | *FRMD4A* | 4 | 0 | Intron | 6 | | - |
| rs7909670 | 3x10$^{-09}$ | Height | *CCDC3* | *CAMK1D* | 4 | 0 | Intergenic | 0 | | - |
| rs12355688 | 6x10$^{-06}$ | Breast cancer | *ZMIZ1* | *ZMIZ1* | 5 | 0 | Non coding transcript exon variant | 4 | | *USF2* |
| rs2145998 | 4x10$^{-13}$ | Height | *PPIF* | *ZMIZ1* | 5 | 0 | Intergenic | 5 | | - |
| rs704010 | 4x10$^{-09}$ | Breast cancer | *ZMIZ1* | *ZMIZ1* | 5 | 0 | Intron | 2b | | *CTCF* |
| rs780151 | 2x10$^{-09}$ | Height | *ZMIZ1* | *ZMIZ1* | 5 | 1 | Intron | 5 | | - |
| rs7916441[a] | 6x10$^{-10}$ | Height | *ZMIZ1* | *ZMIZ1* | 5 | 1 | Intron | 5 | | - |
| rs941873 | 4x10$^{-07}$ | Height | *ZCCHC24* | *ZMIZ1* | 5 | 0 | Intron | 4 | | *HSF1* |
| rs10510102 | 2x10$^{-06}$ | Breast cancer | *ATE1* | *NSMCE4A* | 6 | 0 | Intron | 1f | *ATE1* | - |
| rs10510126 | 7x10$^{-07}$ | Breast cancer | *BUB3* | *HMX2* | 6 | 0 | Intergenic | 5 | | *CREBBP* |
| rs10736303 | 9x10$^{-06}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | | - |
| rs1078806 | 2x10$^{-06}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | | - |
| rs11199914 | 2x10$^{-08}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 0 | Intergenic | 0 | | - |
| rs11200014 | 1x10$^{-05}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | | - |
| rs1219642 | 3x10$^{-16}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 0 | Intron | 6 | | - |
| rs1219648 | 1x10$^{-10}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 0 | | - |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs2420946 | $2x10^{-10}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 0 | - |
| rs2912774 | $3x10^{-15}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | - |
| rs2936870 | $5x10^{-15}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | - |
| rs2981575 | $1x10^{-08}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 6 | - |
| rs2981578 | $1x10^{-15}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 2b | Oct-1 |
| rs2981579 | $2x10^{-10}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | - |
| rs2981582 | $2x10^{-76}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 5 | - |
| rs3135718 | $7x10^{-05}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 0 | Intron | 5 | - |
| rs3750817 | $8x10^{-08}$ | Breast cancer | *FGFR2* | *FGFR2* | 6 | 1 | Intron | 4 | *E2F1* |
| rs4752571 | $1x10^{-05}$ | Breast cancer | *FGFR2* | *NA* | 6 | 0 | Intron | 0 | - |
| rs6585827 | $2x10^{-06}$ | Height | *PLEKHA1* | *HTRA1* | 6 | 0 | Intron | 5 | *RFX3* |
| rs2107425 | $7x10^{-06}$ | Breast cancer | *MRPL23* | *MRPL23* | 7 | 0 | Intron | 4 | *CTCF* |
| rs2237886 | $2x10^{-13}$ | Height | *KCNQ1* | *KCNQ1* | 7 | 0 | Intron | 4 | *USF1* |
| rs3817198 | $2x10^{-11}$ | Breast cancer | *LSP1* | *LSP1* | 7 | 0 | Intron | 5 | - |
| rs909116 | $7x10^{-07}$ | Breast cancer | *TNNT3* | *LSP1* | 7 | 0 | Intron | 5 | - |
| rs3782089 | $4x10^{-13}$ | Height | *SSSCA1* | *OVOL1* | 8 | 0 | Non coding transcript exon variant | 5 | - |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs3903072 | $9x10^{-12}$ | Breast cancer | *SNX32* | *EFEMP2* | 8 | 0 | Intergenic | 5 | - |
| rs3824999 | $4x10^{-10}$ | Colorectal cancer | *POLD3* | *CHRDL2* | 9 | 0 | Intron | 2b | *AP-3* |
| rs606452 | $2x10^{-09}$ | Height | *SERPINH1* | *SERPINH1* | 9 | 1 | Intron | 5 | - |
| rs634552 | $4x10^{-13}$ | Height | *SERPINH1* | *SERPINH1* | 9 | 1 | Intron | 5 | - |
| rs11820646 | $1x10^{-09}$ | Breast cancer | *BARX2* | *BARX2* | 10 | 0 | Regulatory region variant | 4 | *USF1* |
| rs654723 | $4x10^{-11}$ | Height | *FLI1* | *FLI1* | 10 | 0 | Intron | 5 | - |
| rs7107217 | $5x10^{-07}$ | Breast cancer | *BARX2* | *BARX2* | 10 | 0 | Intergenic | 5 | - |
| rs10771399 | $2x10^{-12}$ | Breast cancer | *PTHLH* | *PTHLH* | 11 | 0 | Intergenic | 0 | - |
| rs10843164 | $6x10^{-12}$ | Height | *CCDC91* | *CCDC91* | 11 | 0 | Intron | 0 | - |
| rs2638953 | $7x10^{17}$ | Height | *CCDC91* | *CCDC91* | 11 | 0 | Intron | 0 | - |
| rs7313833 | $6x10^{-06}$ | Breast cancer | *PTHLH* | *PTHLH* | 11 | 0 | Intergenic | 0 | - |
| rs10492321 | $7x10^{-11}$ | Height | *SOCS2* | *SOCS2* | 12 | 1 | Downstream gene variant | 6 | - |
| rs10859563 | $3x10^{-12}$ | Height | *CRADD* | *CRADD* | 12 | 0 | Intron | 0 | - |
| rs11107116 | $1x10^{-34}$ | Height | *SOCS2* | *SOCS2* | 12 | 1 | Downstream gene variant | 6 | - |
| rs3825199 | $2x10^{-07}$ | Height | *SOCS2* | *SOCS2* | 12 | 1 | 3 prime UTR variant | 0 | - |
| rs4761470 | $6x10^{-07}$ | Breast cancer | *PLXNC1* | *PLXNC1* | 12 | 0 | Intron | 5 | - |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs11571833 | $5x10^{-08}$ | Breast cancer | *BRCA2* | *BRCA2* | 13 | 0 | Stop gained | 6 | - |
| rs718444 | $2x10^{-10}$ | Height | *PDS5B* | *N4BP2L1* | 13 | 1 | Intergenic | 6 | - |
| rs7332115 | $6x10^{-10}$ | Height | *PDS5B* | *N4BP2L1* | 13 | 1 | Intergenic | 6 | - |
| rs17104630 | $8x10^{-06}$ | Height | *NKX2-8* | *NKX2-8* | 14 | 0 | Downstream gene variant | 0 | - |
| rs2236007 | $2x10^{-13}$ | Breast cancer | *PAX9* | *PAX9* | 14 | 0 | Intron | 5 | - |
| rs1570106 | $8x10^{-09}$ | Height | *RAD51B* | *RAD51B* | 15 | 0 | Intron | 0 | - |
| rs2588809 | $1x10^{-10}$ | Breast cancer | *RAD51B* | *RAD51B* | 15 | 0 | Intron | 0 | - |
| rs999737 | $2x10^{-07}$ | Breast cancer | *RAD51B* | *RAD51B* | 15 | 0 | Intron | 6 | - |
| rs7153027 | $1x10^{-10}$ | Height | *TRIP11* | *FBLN5* | 16 | 1 | Intergenic | 6 | - |
| rs7155279 | $1x10^{-10}$ | Height | *TRIP11* | *FBLN5* | 16 | 1 | Intron | 6 | - |
| rs8007661 | $6x10^{-10}$ | Height | *TRIP11* | *FBLN5* | 16 | 1 | Intron | 0 | - |
| rs941764 | $4x10^{-10}$ | Breast cancer | *CCDC88C* | *SMEK1* | 16 | 0 | Intron | 4 | *NR3C1, CREBBP* |
| rs3884558 | $4x10^{-06}$ | Breast cancer | *RORA* | *RORA* | 17 | 0 | Intron | 0 | - |
| rs7178424 | $6x10^{-09}$ | Height | *C2CD4A* | *FAM148A* | 17 | 0 | Upstream gene variant | 2b | *STAT3, MAFK* |
| rs2072153 | $4x10^{-08}$ | Height | *ZNF652* | *ZNF652* | 18 | 0 | Intron | 6 | - |
| rs2075555 | $8x10^{-08}$ | Breast cancer | *COL1A1* | *COL1A1* | 18 | 0 | Intron | 4 | *POLR2A, CEBPB, JUN* |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs4605213 | $3x10^{-08}$ | Height | *MBTD1* | *NME2* | 18 | 0 | Intron | 2b | *p300, CTCF, HDAC2, ZNF263* |
| rs12449568 | $2x10^{-06}$ | Height | *ANKFN1* | *ANKFN1* | 19 | 0 | Intron | 0 | - |
| rs1549519 | $6x10^{-09}$ | Height | *TMEM100* | *TMEM100* | 19 | 0 | Intergenic | 6 | - |
| rs227724 | $7x10^{-15}$ | Height | *C17orf67* | *NOG* | 19 | 0 | Intergenic | 0 | - |
| rs4794665 | $1x10^{-07}$ | Height | *C17orf67* | *C17orf67* | 19 | 0 | Intergenic | 3a | *AIRE, GATA1, Foxa2* |
| rs6504950 | $2x10^{-13}$ | Breast cancer | *STXBP4* | *COX11* | 19 | 0 | Intron | 0 | - |
| rs11082671 | $2x10^{-08}$ | Height | *CTIF* | *NA* | 20 | 0 | Intron | 0 | - |
| rs12953717 | $9x10^{-12}$ | Colorectal cancer | *SMAD7* | *SMAD7* | 20 | 1 | Intron | 5 | - |
| rs1787200 | $1x10^{-10}$ | Height | *DYM* | *RPL17* | 20 | 0 | Intron | 6 | *Foxl1, Fokl1, Foxa2, Freac-3* |
| rs4464148 | $3x10^{-08}$ | Colorectal cancer | *SMAD7* | *SMAD7* | 20 | 0 | Intron | 4 | *EBF1* |
| rs4939827 | $1x10^{-07}$ | Colorectal cancer | *SMAD7* | *SMAD7* | 20 | 1 | Intron | 5 | *Elf-1, ELF3, PU.1, Spic* |
| rs8099594 | $3x10^{-07}$ | Height | *DYM* | *RPL17* | 20 | 0 | Upstream gene variant | 3a | *CTCF, TP53* |
| rs9967417 | $5x10^{-09}$ | Height | *DYM* | *RPL17* | 20 | 0 | Intron | 0 | - |
| rs2279008 | $3x10^{-08}$ | Height | *MYO9B* | *MYO9B* | 21 | 0 | Intron | 5 | - |
| rs8100241 | $4x10^{-08}$ | Breast cancer | *ANKLE1* | *C19orf62* | 21 | 0 | Missense variant | 4 | *GATA1, TAF1* |
| rs8170 | $2x10^{-09}$ | Breast cancer | *BABAM1* | *C19orf62* | 21 | 0 | Synonymous variant | 4 | *CEBPB* |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| rs10187066 | $2x10^{-07}$ | Height | *ZNF142* | *STK36* | 22 | 0 | Intron | 1f | *SLC11A1, CYP27A1* | |
| rs1052483 | $1x10^{-06}$ | Height | *NHEJ1* | *IHH* | 22 | 1 | Non coding transcript exon variant | 1f | *SLC23A3* | - |
| rs12470505 | $9x10^{-12}$ | Height | *CCDC108/IHH* | *IHH* | 22 | 1 | Upstream gene variant | 1f | *SLC23A3* | *USF1* |
| rs13387042 | $1x10^{-13}$ | Breast cancer | *TNP1* | *TNP1* | 22 | 0 | Intergenic | 0 | | - |
| rs1351164 | $2x10^{-14}$ | Height | *TNS1* | *NA* | 22 | 0 | Intron | 0 | | - |
| rs16857609 | $1x10^{-15}$ | Breast cancer | *TNS1* | *NA* | 22 | 0 | Intron | 5 | | - |
| rs16859517 | $5x10^{-06}$ | Height | *SLC23A3* | *IHH* | 22 | 0 | Intergenic | 0 | | - |
| rs2553026 | $6x10^{-08}$ | Height | *TNP1* | *NA* | 22 | 0 | Intron | 0 | | - |
| rs3791950 | $2x10^{-06}$ | Height | *TNS1* | *TNS1* | 22 | 0 | Intron | 2b | | *PAX5* |
| rs6435999 | $7x10^{-07}$ | Height | *TNS1* | *DIRC3* | 22 | 0 | Intergenic | 0 | | - |
| rs6724465 | $2x10^{-08}$ | Height | *SLC23A3 / NHEJ1* | *IHH* | 22 | 1 | Intron | 1f | *SLC23A3* | - |
| rs2145270 | $5x10^{-18}$ | Height | *BMP2* | *BMP2* | 23 | 1 | Intergenic | 0 | | - |
| rs2145272 | $2x10^{-24}$ | Height | *BMP2* | *BMP2* | 23 | 1 | Intergenic | 5 | | *NFKB1, BATF* |
| rs4813802 | $7x10^{-06}$ | Colorectal cancer | *BMP2* | *BMP2* | 23 | 0 | promotor flanking region | 4 | | *STAT3, TRIM28, TFAP2A, TFAP2C, USF2, SETDB1, HNF4a* |
| rs961253 | $2x10^{-10}$ | Colorectal cancer | *FERMT1* | *FERMT1* | 23 | 0 | Intergenic | 5 | | - |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs967417 | $2x10^{-08}$ | Height | *BMP2* | *BMP2* | 23 | 1 | Intergenic | 6 | - |
| rs1074683 | $1x10^{-14}$ | Height | *PXMP4* | *CBFA2T2* | 24 | 1 | Intron | 0 | - |
| rs2284378 | $1x10^{-08}$ | Breast cancer | *RALY* | *RALY* | 24 | 0 | Intron | 0 | - |
| rs7274811 | $6x10^{-22}$ | Height | *ZNF341* | *CBFA2T2* | 24 | 1 | Intron | 0 | - |
| rs139909 | $2x10^{-07}$ | Height | *TNRC6B* | *TNRC6B* | 25 | 0 | Intron | 2b | *BATF* |
| rs5757949 | $4x10^{-06}$ | Height | *MKL1* | *MCHR1* | 25 | 0 | Intron | 5 | - |
| rs6001930 | $2x10^{-06}$ | Breast cancer | *MKL1* | *MCHR1* | 25 | 0 | Intron | 0 | - |
| rs4552313 | $2x10^{-06}$ | Height | *CMC1* | *ZCWPW2* | 26 | 0 | Intron | 6 | - |
| rs4973768 | $2x10^{-08}$ | Breast cancer | *SLC4A7* | *NEK10* | 26 | 0 | 3 prime UTR variant | 6 | - |
| rs10010325 | $4x10^{-11}$ | Height | *TET2* | *TET2* | 27 | 0 | Intron | 0 | - |
| rs6855629 | $2x10^{-08}$ | Height | *TET2* | *EEF1AL7* | 27 | 0 | Intron | 6 | - |
| rs9790517 | $4x10^{-08}$ | Breast cancer | *TET2* | *TET2* | 27 | 0 | Intron | 0 | - |
| rs13177718 | $3x10^{-08}$ | Height | *FER* | *FER* | 28 | 0 | Intron | 6 | - |
| rs367615 | $4x10^{-08}$ | Colorectal cancer | *MAN2A1* | *PJA2* | 28 | 0 | Intergenic | 5 | - |
| rs31198 | $8x10^{-06}$ | Height | *PITX1* | *PITX1* | 29 | 0 | Intron | 5 | - |
| rs526896 | $2x10^{-13}$ | Height | *PITX1* | *PITX1* | 29 | 0 | Intergenic | 5 | - |

The table Continues on the next page

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| rs647161 | $1x10^{-10}$ | Colorectal cancer | *PITX1* | *PITX1* | 29 | 0 | Intron | 6 | - |
| rs6879260 | $2x10^{-09}$ | Height | *GFPT2* | *GFPT2* | 30 | 0 | Intron | 4 | *USF1* |
| rs7711990 | $8x10^{-05}$ | Breast cancer | *BTNL8* | *NA* | 30 | 0 | Non coding transcript exon variant | 6 | - |
| rs204247 | $8x10^{-09}$ | Breast cancer | *RANBP9* | *RANBP9* | 31 | 0 | Intergenic | 6 | - |
| rs853356 | $3x10^{-06}$ | Height | *CD83* | *RNF182* | 31 | 0 | Intergenic | 0 | - |
| rs1047014 | $2x10^{-13}$ | Height | *ID4* | *ID4* | 32 | 0 | Upstream gene | 5 | - |
| rs16882214 | $2x10^{-06}$ | Breast cancer | *ID4* | *NA* | 32 | 0 | Intergenic | 0 | - |
| rs17530068 | $3x10^{-06}$ | Breast cancer | *FAM46A* | *FAM46A* | 33 | 0 | Intergenic | 0 | - |
| rs2322633 | $3x10^{-09}$ | Height | *BCKDHB* | *BCKDHB* | 33 | 0 | Intron | 6 | - |
| rs310405 | $1x10^{-10}$ | Height | *FAM46A* | *NA* | 33 | 0 | Intergenic | 0 | - |
| rs2057314 | $3x10^{-06}$ | Colorectal cancer | *DCBLD1* | *DCBLD1* | 34 | 0 | Intron | 4 | *SPI1* |
| rs9285425 | $2x10^{-08}$ | Height | *DCBLD1* | *DCBLD1* | 34 | 0 | Intron | 0 | - |
| rs961764 | $1x10^{-11}$ | Height | *VGLL2* | *RFXDC1* | 34 | 0 | Intergenic | 0 | - |
| rs1361108 | $9x10^{-06}$ | Height | *CENPW* | *C6orf173* | 35 | 1 | Intergenic | 5 | - |
| rs1490384 | $1x10^{-16}$ | Height | *CENPW* | *C6orf173* | 35 | 1 | Intergenic | 0 | - |
| rs1490388 | $6x10^{-07}$ | Height | *CENPW* | *C6orf173* | 35 | 1 | Intergenic | 0 | - |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| rs2180341 | $3x10^{-08}$ | Breast cancer | *RNF146* | *ECHDC1* | 35 | 0 | Intron | 0 | | - |
| rs4549631 | $5x10^{-13}$ | Height | *CENPW* | *C6orf173* | 35 | 1 | Downstream gene variant | 0 | | - |
| rs2046210 | $2x10^{-15}$ | Breast cancer | *CCDC170* | *C6orf97* | 36 | 0 | Intergenic | 1f | *C6orf97* | - |
| rs2982712 | $4x10^{-10}$ | Height | *ESR1* | *ESR1* | 36 | 0 | Intron | 0 | | - |
| rs3734805 | $1x10^{-07}$ | Breast cancer | *CCDC170* | *C6orf97* | 36 | 1 | Intron | 0 | | - |
| rs3757318 | $2x10^{-21}$ | Breast cancer | *CCDC170* | *C6orf97* | 36 | 1 | Intron | 4 | | *HNF4A, HNF4G* |
| rs543650 | $1x10^{-17}$ | Height | *ESR1* | *ESR1* | 36 | 0 | Intron | 0 | | - |
| rs9383938 | $2x10^{-10}$ | Breast cancer | *ESR1* | *C6orf97* | 36 | 0 | Intron | 5 | | *RFX3* |
| rs9383951 | $2x10^{-06}$ | Breast cancer | *ESR1* | *ESR1* | 36 | 0 | Intron | 4 | | *GATA2* |
| rs9365723 | $4x10^{-06}$ | Colorectal cancer | *SYNJ2* | *SYNJ2* | 37 | 0 | Intron | 5 | | - |
| rs9456307 | $2x10^{-09}$ | Height | *TULP4* | *GTF2H5* | 37 | 0 | 3 prime UTR variant | 0 | | - |
| rs2128382 | $8x10^{-06}$ | Colorectal cancer | *GSDMC* | *FAM49B* | 38 | 0 | Intergenic | 0 | | - |
| rs6470764 | $2x10^{-28}$ | Height | *GSDMC* | *MLZE* | 38 | 0 | Enhancer region | 5 | | - |
| rs10114408 | $3x10^{-06}$ | Colorectal cancer | *BARX1* | *BARX1* | 39 | 0 | Intergenic | 6 | | - |
| rs10512248 | $4x10^{-11}$ | Height | *PTCH1* | *PTCH1* | 39 | 0 | Intron | 6 | | - |
| rs10816533 | $2x10^{-06}$ | Height | *ZNF510* | *ZNF782* | 39 | 0 | Intron | 1f | *LOC642921* | |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| rs1257763 | $1x10^{-09}$ | Height | *PTPDC1* | *BARX1* | 39 | 0 | Intergenic | 0 | | - |
| rs16910061 | $3x10^{-06}$ | Height | *FBP2* | *FBP1* | 39 | 0 | Upstream gene variant | 5 | | *JUND* |
| rs2025151 | $2x10^{-12}$ | Height | *ZNF367* | *HABP4* | 39 | 0 | Intron | 1f | *HABP4* | *POLR2A* |
| rs473902 | $2x10^{-17}$ | Height | *PTCH1* | *PTCH1* | 39 | 0 | Intron | 5 | | - |
| rs10759243 | $1x10^{-08}$ | Breast cancer | *KLF4* | *KLF4* | 40 | 0 | Upstream gene variant | 0 | | - |
| rs4743034 | $2x10^{-08}$ | Height | *ZNF462* | *ZNF462* | 40 | 0 | Intron | 5 | | - |
| rs7027110 | $2x10^{-13}$ | Height | *ZNF462* | *ZNF462* | 40 | 0 | Intergenic | 0 | | - |
| rs865686 | $1x10^{-34}$ | Breast cancer | *KLF4* | NA | 40 | 0 | Intergenic | 0 | | - |

Abbreviations: eQTL; expression Quantitative Trait Locus GRAIL, Gene Relationships Among Implicated Loci; GWAS, genome-wide association study; ID, identification number; LD, linkage disequilibrium; NA, not available in GWAS catalogue; SNP, single nucleotide polymorphism. References: 1. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009;106(23):9362-7. doi: 10.1073/pnas.0903103106. PubMed PMID: 19474294; PubMed Central PMCID: PMC2687147. 2. Johnson AD, O'Donnell CJ. An open access database of genome-wide association results. BMC medical genetics. 2009;10:6. doi: 10.1186/1471-2350-10-6. PubMed PMID: 19161620; PubMed Central PMCID: PMC2639349.

[a] Gene annotations were based on the physical mapping of a SNP according to HapMap.

[b] Gene annotations using GRAIL were based on gene relationships among the complete set of SNPs listed in this table.

[c] An LD tag equal to one denotes that two or more SNPs within the same cluster are in high LD ($r^2 > 0.7$). For the prioritised clusters, SNPs with the highest level of regulatory evidence were prioritised. In cases where the regulatory evidence was equal, SNPs were prioritised according to the most significant p-value for the tested association.

[d] The score denotes the scale from RegulomeDB, with scores 1a-1f denoting SNPs that were likely linked to the expression of a gene target (cis-eQTLs), scores 2-3 denoting SNPs that likely only affected protein binding, scores 4-6 denoting SNPs for which there was minimal binding evidence, and score 0 denoting SNPs for which no evidence was available.

[e] Gene target of cis-eQTL (cis-eQTLs are SNPs that are associated with the regulation of mRNA expression of a nearby located gene) .

[f] Evidence for transcription factor binding at the genomic coordinate of a SNP according to RegulomeDB. The number in brackets on the left hand-side of a transcription factor indicates the number of transcription factor binding proteins that are binding to the genomic coordinate of a SNP.

**Table S2. Characteristics of SNPs within clusters that included at least one height-associated GWAS SNP and one post-menopausal breast or colorectal cancer risk-associated GWAS SNP annotated to the same gene based on either HapMap or GRAIL**

| Cluster ID | SNP ID[a] | Gene annotation[b] | First author GWAS | Phenotype[c] | P-value | OR for breast or colorectal cancer, or beta-coefficient for each unit increase or decrease in height [95% CI] | Ancestry[d] | Risk allele[e] | MAF non-CEU[f] | MAF CEU[g] |
|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster ID 22** | rs13387042 | *TNP1* | Michailidou*et al.*, 2013 | Breast cancer | $2x10^{-57}$ | 1.14 [1.11-1.16] | European | 0.51 (A) | | 0.43 (G) |
| | | | Fletcher *et al.*, 2011 | Breast cancer | $2x10^{-10}$ | 1.16 [1.11-1.22] | European | 0.52 (A) | | 0.43 (G) |
| | | | Li *et al.*, 2010 | Breast cancer | $9x10^{-06}$ | 1.18 [1.10-1.27] | European | 0.53 (A) | | 0.43 (G) |
| | | | Turnbull *et al.*, 2010 | Breast cancer | $2x10^{-10}$ | 1.21 [1.14-1.29] | European | 0.49 (A) | | 0.43 (G) |
| | | | Thomas *et al.*, 2009 | Breast cancer | $2x10^{-08}$ | 1.25 (Het) [1.15-1.37] | European | 0.51 (A) | | 0.43 (G) |
| | | | Stacey *et al.*, 2007 | Breast cancer | $1x10^{-13}$ | 1.2 [1.14-1.26] | European | 0.50 (A) | | 0.43 (G) |
| | rs2553026 | *TNP1* | N'Diaye *et al.*, 2011 | Height | $6x10^{-08}$ | 0.056 [0.036-0.076] unit increase | African | 0.19 (A) | 0.23 (A) | 0.30 (G) |
| | rs1351164 | *TNS1* | Lango Allen *et al.*, 2010 | Height | $2x10^{-14}$ | 0.034 [NA] unit increase | European | 0.79 (T) | | 0.19 (C) |
| | rs16857609 | *TNS1* | Michailidou*et al.*, 2013 | Breast cancer | $1x10^{-15}$ | 1.08 [1.06-1.10] | European | 0.26 (T) | | 0.29 (T) |
| | rs6435999 | *TNS1* | N'Diaye *et al.*, 2011 | Height | $7x10^{-07}$ | 0.041 [0.025-0.057] unit increase | African | 0.64 (A) | 0.30 (G) | 0.01 (G) |
| | rs3791950 | *TNS1* | N'Diaye *et al.*, 2011 | Height | $2x10^{-06}$ | 0.061 [0.036-0.086] unit decrease | African | 0.89 (A) | 0.17 (C) | 0.46 (C) |
| | rs10187066 | *ZNF142* | Lango Allen *et al.*, 2010 | Height | $2x10^{-07}$ | NA | European | NA | | 0.36 (A) |
| | rs12470505*[a] | *CCDC108/IHH* | Lango *et al.*, 2010 | Height | $9x10^{-12}$ | 0.041 [NR] unit increase | European | 0.90 (T) | | 0.09 (G) |
| | rs1052483 | *NHEJ1* | Gudbjartsson *et al.*, 2008 | Height | $1x10^{-06}$ | 6.9 [4.16-9.64] % SD taller | European | 0.91 (G) | | 0.09 (T) |
| | rs6724465* | *SLC23A3/NHEJ1* | Weedon *et al.*, 2008 | Height | $2x10^{-08}$ | 0.06 [0.02-0.10] SD shorter - among males | European | 0.10 (A) | | 0.09 (A) |
| | rs16859517* | *SLC23A3* | Okada *et al.*, 2010 | Height | $5x10^{-06}$ | NA | European | NA | | 0.04 (T) |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster ID 27** | rs9790517 | *TET2* | Michailidou*et al.*, 2013 | Breast cancer | $4x10^{-08}$ | 1.05 [1.03-1.08] | European | 0.23 (T) | | 0.21 (T) |
| | rs10010325 | *TET2* | Lango Allen *et al.*, 2010 | Height | $4x10^{-11}$ | 0.024 [NA] unit increase | European | 0.49 (A) | | 0.48 (A) |
| | rs6855629 | *TET2* | Berndt *et al.*, 2013 | Height | $2x10^{-08}$ | 1.14 [NA] | European | 0.63 (G) | | 0.39 (A) |
| **Cluster ID 29** | rs526896*[a] | *PITX1* | Lango Allen *et al.*, 2010 | Height | $2x10^{-13}$ | 1.15 [NA] | European | 0.73 (T) | | 0.29 (G) |
| | | | Berndt *et al.*, 2013 | Height | $9x10^{-10}$ | 0.03 [NA] unit increase | European | 0.72 (T) | | 0.29 (G) |
| | rs31198* | *PITX1* | Gudbjartsson *et al.*, 2008 | Height | $8x10^{-06}$ | 4.8 [2.64-6.96] % SD taller | European | 0.75 (T) | | 0.26 (C) |
| | rs647161 | PITX1 | Jia *et al.*, 2012 | Colorectal cancer | $1x10^{-10}$ | 1.11 [1.08-1.15] | European | 0.67 (A) | | 0.32 (C) |
| | | | Jia *et al.*, 2012 | Colorectal cancer | $4x10^{-10}$ | 1.17 [1.11-1.22] | East-Asian | 0.31 (A) | 0.30 (A) | 0.32 (C) |
| **Cluster ID 32** | rs1047014 | ID4 | Lango Allen *et al.*, 2010 | Height | $2x10^{-13}$ | 0.032 [NA] unit decrease | European | 0.75 (T) | | 0.28 (C) |
| | rs16882214 | *ID4* | Rinella *et al.*, 2013 | Breast cancer | $2x10^{-06}$ | 1.43 [NA] | European | 0.81 (NA) | | 0.14 (G) |
| **Cluster ID 33** | rs2322633 | *BCKDHB* | Berndt *et al.*, 2013 | Height | $3x10^{-09}$ | 1.12 [NA] | European | 0.50 (T) | | 0.48 (C) |
| | rs310405 | *FAM46A* | Berndt *et al.*, 2013 | Height | $1x10^{-10}$ | 1.14 [NA] | European | 0.52 (A) | | 0.47 (G) |
| | | | Lango Allen *et al.*, 2010 | Height | $2x10^{-13}$ | 0.026 [NA] unit increase | European | 0.52 (A) | | 0.47 (G) |
| | rs17530068 | *FAM46A* | Garcia-Closas *et al.*, 2013 | Breast cancer | $3x10^{-06}$ | 1.09 [1.05-1.13] | European | 0.24 (C) | | 0.20 (C) |
| | | | Michailidou*et al.*, 2013 | Breast cancer | $8x10^{-09}$ | 1.05 [1.03-1.08] | European | 0.22 (G) | | 0.20 (C) |
| | | | Siddiq *et al.*, 2012 | Breast cancer | $3x10^{-07}$ | 1.16 [1.10-1.23] | Mixed | 0.24 (C) | 0.18 (C) | 0.20 (C) |
| **Cluster ID 34** | rs961764 | *VGLL2* | Lango Allen *et al.*, 2010 | Height | $1x10^{-11}$ | 0.024 [NA] unit decrease | European | 0.42 (C) | | 0.41 (C) |
| | rs2057314 | *DCBLD1* | Peters *et al.*, 2012 | Colorectal cancer | $3x10^{-06}$ | 1.08 [1.04-1.11] | European | 0.496 (G) | | 0.49 (A) |
| | rs9285425 | *DCBLD1* | Berndt *et al.*, 2013 | Height | $2x10^{-08}$ | 1.14 [NA] | European | 0.50 (G) | | 0.49 (G) |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster ID 7** | rs3757318*[a] | *C6orf97* | Purrington *et al.*, 2013 | Breast cancer | $9x10^{-06}$ | 1.33 [1.17-1.51] | European | NA | | 0.08 (A) |
| | | | Michailidou*et al.*, 2013 | Breast cancer | $2x10^{-21}$ | 1.16 [1.12-1.21] | European | 0.07 (A) | | 0.08 (A) |
| | | | Turnbull *et al.*, 2010 | Breast cancer | $3x10^{-06}$ | 1.30 [1.17-1.46] | European | 0.07 (A) | | 0.08 (A) |
| | rs3734805* | *C6orf97* | Fletcher *et al., 2011* | Breast cancer | $1x10^{-07}$ | 1.19 [1.11-1.27] | European | 0.08 (C) | | 0.06 (C) |
| | rs2046210 | *ESR1* | Garcia-Closas *et al., 2013* | Breast cancer | $5x10^{-16}$ | 1.15 [1.11-1.19] | European | 0.42 (A) | | 0.29 (A) |
| | | | Couch *et al.*, 2013 | Breast cancer | $5x10^{-09}$ | 1.28 [1.18-1.39] | European | 0.08 (C) | | 0.29 (A) |
| | | | Zheng *et al.*, 2009 | Breast cancer | $2x10^{-15}$ | 1.29 [1.21-1.37] | Asian | 0.37 (A) | 0.36 (A) | 0.29 (A) |
| | rs9383938 | *ESR1* | Siddiq *et al.*, 2012 | Breast cancer | $2x10^{-10}$ | 1.28 [NA] | Mixed | NA (T) | 0.15 (T) | 0.06 (T) |
| | rs543650 | *ESR1* | Lango Allen *et al.*, 2010 | Height | $1x10^{-17}$ | 0.034 [NA] unit decrease | European | 0.40 (T) | | 0.39 (T) |
| | rs9383951 | *ESR1* | Long *et al.*, 2012 | Breast cancer | $2x10^{-6}$ | 1.14 [1.08-1.19] | Asian | 0.90 (G) | 0.07 (C) | 0.01 (C) |
| | rs2982712 | *ESR1* | Berndt *et al.*, 2013 | Height | $4x10^{-10}$ | 1.17 [NA] | European | 0.47 (C) | | 0.47 (C) |
| **Cluster ID 39** | rs10114408 | *BARX1* | Jiao *et al.*, 2012 | Colorectal cancer | $3x10^{-06}$ | 1.37 [1.20-1.56] | European | 0.76 (NA) | | 0.26 (T) |
| | rs1257763 | BARX1 | Lango Allen *et al.*, 2010 | Height | $1x10^{-09}$ | 0.069 [NA] unit increase | European | 0.04 (A) | | 0.03 (A) |
| | rs16910061 | *FBP1* | Kim *et al.*, 2009 | Height | $3x10^{-06}$ | 0.53 NA cm decrease | Korean | 0.14 (T) | 0.12 (A) | 0.04 (A) |
| | rs473902 | *PTCH1* | Lango Allen *et al.*, 2010 | Height | $2x10^{-17}$ | 0.069 [NA] unit increase | European | 0.92 (T) | | 0.09 (G) |
| | rs10512248 | *PTCH1* | Weedon *et al.*, 2008 | Height | $4x10^{-11}$ | 0.05 [0.02-0.07] SD taller - among males | European | 0.31 (G) | | 0.33 (G) |
| | rs2025151 | *HABP4* | Berndt *et al.*, 2013 | Height | $2x10^{-12}$ | 1.22 [NA] | European | 0.18 (G) | | 0.17 (G) |
| | rs10816533 | *ZNF782* | Lei *et al.*, 2008 | Height | $2x10^{-06}$ | NA | Chinese | 0.29 (C) | 0.29 (C) | 0.03 (C) |
| **Cluster ID 5** | rs704010 | *ZMIZ1* | Michailidou *et al.*, 2013 | Breast cancer | $7x10^{-22}$ | 1.08 [1.06-1.10] | European | 0.38 (T) | | 0.44 (T) |
| | | | Turnbull *et al.*, 2010 | Breast cancer | $4x10^{-09}$ | 1.07 [1.03-1.11] | European | 0.39 (A) | | 0.44 (T) |
| | rs7916441* | ZMIZ1 | Lango Allen *et al.*, 2010 | Height | $6x10^{-10}$ | NA | European | NA | | 0.48 (C) |

The table Continues on the next page

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | rs780151*[a] | *ZMIZ1* | Berndt *et al.*, 2013 | Height | $2x10^{-09}$ | 1.13 [NA] | European | 0.57 (G) | | 0.46 (A) |
| | rs12355688 | *ZMIZ1* | Song *et al.*, 2013 | Breast cancer | $6x10^{-06}$ | 1.24 [1.13-1.36] | African | 0.22 (T) | 0.21 (T) | 0.05 (T) |
| | rs2145998* | *PPIF* | Lango Allen *et al.*, 2010 | Height | $4x10^{-13}$ | 0.026 [NA] unit decrease | European | 0.49 (A) | | 0.48 (A) |
| | rs941873*[a] | *ZCCHC24* | N'Diaye *et al.*, 2011 | Height | $4x10^{-07}$ | NA | African | 0.41 (A) | 0.42 (A) | 0.49 (A) |
| **Cluster ID 15** | rs2588809 | RAD51B | Michailidou*et al.*, 2013 | Breast cancer | $1x10^{-10}$ | 1.08 [1.05-1.11] | European | 0.16 (T) | | 0.18 (T) |
| | rs1570106 | *RAD51B* | Lango Allen *et al.*, 2010 | Height | $8x10^{-09}$ | 0.026 [NA] unit decrease | European | 0.20 (T) | | 0.21 (T) |
| | rs999737 | *RAD51B* | Michailidou*et al.*, 2013 | Breast cancer | $3x10^{-19}$ | 1.09 [1.06-1.11] | European | 0.77 (C) | | 0.27 (T) |
| | | | Thomas *et al.*, 2009 | Breast cancer | $2x10^{-07}$ | 1.06 (Het) [1.01-1.14] | European | 0.76 (C) | | 0.27 (T) |
| **Cluster ID 23** | rs961253 | *FERMT1* | Houlston *et al.*, 2008 | Colorectal cancer | $2x10^{-10}$ | 1.12 [1.08-1.16] | European | 0.36 (A) | | 0.41 (A) |
| | rs967417* | *BMP2* | Gudbjartsson *et al.*, 2008 | Height | $2x10^{-08}$ | 4.3 [2.73-5.87] % SD taller | European | 0.53 (C) | | 0.40 (A) |
| | rs2145270 | *BMP2* | Berndt *et al.*, 2013 | Height | $5x10^{-18}$ | 1.2 [NA] | European | 0.37 (C) | | 0.42 (C) |
| | rs2145272*[a] | *BMP2* | Lango Allen *et al.*, 2010 | Height | $2x10^{-24}$ | 0.039 [NA] unit decrease | European | 0.65 (A) | | 0.41 (G) |
| | rs4813802 | *BMP2* | Peters *et al.*, 2012 | Colorectal cancer | $7x10^{-06}$ | 1.1 [1.05-1.14] | European | 0.34 (G) | | 0.35 (G) |
| **Cluster ID 25** | rs139909 | *TNRC6B* | Estrada *et al.*, 2009 | Height | $2x10^{-07}$ | 0.25 [0.03-0.47] cm increase | European | 0.68 (T) | | 0.32 (C) |
| | rs5757949 | *MKL1* | Estrada *et al.*, 2009 | Height | $4x10^{-06}$ | NA | European | NA (T) | | 0.31 (C) |
| | rs6001930 | *MKL1* | Garcia-Closas *et al.*, 2013 | Breast cancer | $2x10^{-06}$ | 1.14 [1.08-1.20] | European | 0.11 (C) | | 0.09 (C) |
| | | | Michailidou*et al.*, 2013 | Breast cancer | $9x10^{-19}$ | 1.12 [1.09-1.16] | European | 0.11 (C) | | 0.09 (C) |

The table Continues on the next page

Abbreviations: CEU, Utah residents with Northern and Western European ancestry; CI, confidence interval;GRAIL, Gene Relationships Among Implicated Loci; GWAS, genome-wide association study; Het, heterozygous genotype; ID, identification number; MAF, minor allele frequency; NA, not available in GWAS catalogue; OR, odds ratio; SD, standard deviation; SNP, single nucleotide polymorphism.

[a] SNPs within clusters that are in high LD with each other ($r^2 > 0.7$) are indicated with a star. SNPs with the highest level of regulatory evidence were prioritised and are indicated by the footnote ([a]). In cases where the regulatory evidence was equal, SNPs in high LD were prioritised according to the most significant p-value.

[b] Gene annotations using GRAIL were based on gene relationships among the complete set of SNPs listed in this table.

[c] Phenotype specifies whether the GWAS SNP was associated with height, breast cancer risk or colorectal cancer risk.

[d] Ethnicity of the population in which the GWAS was conducted.

[e] Risk allele frequency in controls as reported in each individual GWAS study.

[f] Minor allele frequency in non-CEU populations for a given GWAS SNP according to data on population genetics in the 1000Genomes project as can be derived from the Ensembl Genome Browser.

[g] Minor allele frequency in CEU populations for a given GWAS SNP according to data on population genetics in the 1000Genomes project as can be derived from the Ensembl Genome Browser.

**Table S3. Overrepresented pathways-using the genes annotated to the SNPs in all identified SNP clusters (i.e. before the prioritisation step in which clusters were prioritised that included at least one height- and one post-menopausal breast or colorectal cancer risk-associated SNP annotated to the same gene)**

| Pathway name[a] | Set size | Number of genes from set in annotated gene list | *p*-value | *q*-value[b] | Pathway source |
|---|---|---|---|---|---|
| Homologous recombination | 28 | 3 | $3.8x10^{-4}$ | $4.6x10^{-2}$ | KEGG |
| Endoderm differentiation | 71 | 4 | $4.7x10^{-4}$ | $4.6x10^{-2}$ | Wikipathways |
| Homologous recombination | 28 | 2 | $1.9x10^{-3}$ | $9.1x10^{-2}$ | Wikipathways |
| Collagen biosynthesis and modifying enzymes | 68 | 3 | $4.1x10^{-3}$ | $9.1x10^{-2}$ | Reactome |
| Endochondral Ossification | 64 | 3 | $4.1x10^{-3}$ | $9.1x10^{-2}$ | Wikipathways |
| Signalling by BMP | 21 | 2 | $4.1x10^{-3}$ | $9.1x10^{-2}$ | Reactome |
| BMP Signalling Pathway | 21 | 2 | $5.1x10^{-3}$ | $9.1x10^{-2}$ | HumanCyc |
| Signalling pathways regulating pluripotency of stem cells | 142 | 4 | $5.8x10^{-3}$ | $9.1x10^{-2}$ | PID |
| Signalling events mediated by the Hedgehog family | 23 | 2 | $6.1x10^{-3}$ | $9.1x10^{-2}$ | PID |
| Regulation of nuclear SMAD2/3 signalling | 77 | 3 | $6.9x10^{-3}$ | $9.1x10^{-2}$ | PID |

Abbreviations: BMP, bone morphogenetic protein; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PID, Pathway Interaction Database; SMAD, a set of protein homologs of both the Drosophila protein, mothers against decapentaplegic (MAD) and the Caenorhabditis elegans protein "SMA" (from gene SMA for small body size); SNP, single nucleotide polymorphism.

[a] Overrrepresented pathways were retrieved using the SNP-gene annotations from GRAIL.

[b] The *p*-values are corrected for multiple testing using the false discovery rate method and are shown as *q*-values.

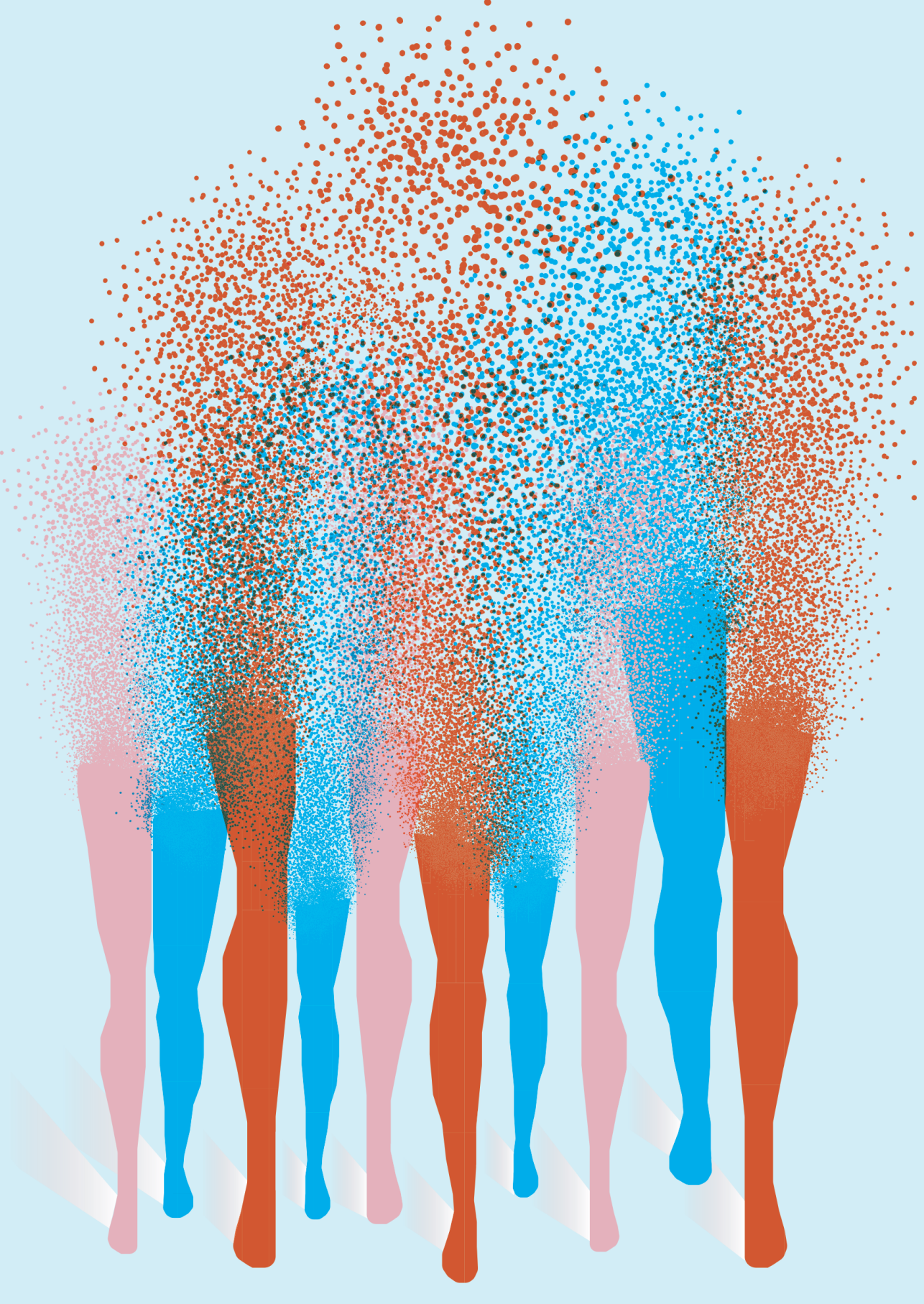
**Table S4. Top ten most significantly overrepresented ontology terms using the genes annotated to the SNPs in all identified SNP clusters (i.e. before the prioritisation step in which clusters were prioritised that included at least one height- and one post-menopausal breast or colorectal cancer risk-associated SNP annotated to the same gene)**

| GO terms[a] | Set size | Number of genes from set in annotated gene list | *p*-value[b] | *q*-value |
|---|---|---|---|---|
| GO:0009887 organ morphogenesis | 918 | 19 | $5.3x10^{-8}$ | $1.3x10^{-5}$ |
| GO:0060348 bone development | 176 | 8 | $2.0x10^{-6}$ | $2.4x10^{-4}$ |
| GO:0001501 skeletal system development | 488 | 12 | $3.4x10^{-6}$ | $1.9x10^{-3}$ |
| GO:0048732 gland development | 407 | 11 | $3.4x10^{-6}$ | $2.7x10^{-4}$ |
| GO:0009653 anatomical structure morphogenesis | 2484 | 28 | $8.0x10^{-6}$ | $4.8x10^{-4}$ |
| GO:0090304 nucleic acid metabolic process | 4893 | 42 | $1.9x10^{-5}$ | $2.1x10^{-3}$ |
| GO:0010467 gene expression | 5291 | 44 | $2.3x10^{-5}$ | $2.1x10^{-3}$ |
| GO:0060255 regulation of macromolecule metabolic process | 5358 | 44 | $3.9x10^{-5}$ | $2.1x10^{-3}$ |
| GO:0043433 negative regulation of sequence-specific DNA binding transcription factor activity | 128 | 6 | $3.3x10^{-5}$ | $2.1x10^{-3}$ |
| GO:0030154 cell differentiation | 3504 | 33 | $3.9x10^{-5}$ | $2.1x10^{-3}$ |

Abbreviations: GO, gene ontology; SNP, single nucleotide polymorphism.

[a] Overrepresentation analysis for GO terms were performed using using the SNP-gene annotations from GRAIL.

[b] The *p*-values are corrected for multiple testing using the false discovery rate method and are available as *q*-values.



Chapter 5

# General discussion

# 1. Summary of main findings

In Chapter 2, a systematic review and meta-analysis was conducted to systematically review the human observational literature on early life energy restriction and cancer risk and to provide quantitative pooled estimates of site-specific cancer risk. The review indicated that moderate continuous energy restriction, as studied in animal experimental and human ecological studies, was generally associated with a decreased site-specific cancer risk. The meta-analysis of human observational studies indicated that severe transient early life energy restriction was associated with a 28% increased breast cancer risk and a 16% increased prostate cancer risk, though some of the underlying studies showed null results. The evidence for an association between severe transient early life energy restriction and risk at other cancer sites, *i.e.* colorectal-, stomach-, pancreas-, ovarian-, and respiratory cancer was either limited or studies were too heterogeneous for pooling. A subsequent meta-regression analysis investigating the effect of the duration and severity of energy restriction on overall cancer risk indicated that duration, rather than severity might result in increased overall cancer risk in women and men, however, this result should be interpreted with caution. With regard to timing of exposure to early life energy restriction, no conclusions could be drawn with regard to breast cancer risk. The sub analysis on energy restriction in women aged 10 to 20 years indicated a 21% increased breast cancer risk compared to women not exposed during that age period. However, a comparison with the summary risk estimate in women aged 0 to 10 years was not possible as this estimate could not be computed due to high between-study heterogeneity.

Chapter 3 describes associations of adult-attained height and early life energy restriction with postmenopausal breast cancer risk according to estrogen and progesterone receptor status within a large prospective cohort study, the Netherlands Cohort Study. Adult-attained height was significantly positively associated with postmenopausal breast cancer risk, in particular with hormone sensitive- subtypes. Of the three exposures to energy restriction investigated, *i.e.* exposure to energy restriction during the the Hunger Winter (the winter of 1944-45), the War Years (1940-44), and the Economic Depression (1932-40), only exposure to energy restriction during the Economic Depression was related to a shorter stature (an almost 2 cm reduction) in female subcohort members. Energy restriction during all three periods of exposure, provided that the exposure occurred before and/or during the growth spurt, was associated with a

significantly decreased risk of hormone receptor-positive breast cancer subtypes. Interestingly, energy restriction during the Hunger Winter increased the estrogen receptor-negative breast cancer risk regardless of the timing of energy restriction. Taken together, the observation that both height and early life energy restriction taking place before and/or during the growth spurt were associated with reduced hormone receptor-positive breast cancer risk seems to suggest possible common underlying mechanisms and critical exposure periods in life.

In Chapter 4, a SNP selection strategy is presented in which SNPs from GWAS repositories are selected for genotyping in large-scale studies investigating shared mechanisms between diseases. The SNP selection strategy is based on the assumption that SNPs from GWAS associated with complex diseases or traits tend to co-segregate in regions of low recombination, harboring functionally linked gene clusters. This phenomenon allows for selecting a limited number of SNPs. This SNP selection approach is of particular interest for studies with exhaustive bio-samples, in which a genome-wide approach is not feasible, and will reduce the costs of genotyping and the chance of false-positive findings. The novelty of this method is the comprehensive integration of publically available GWAS repositories, on the basis of which SNPs associated with multiple associated complex traits and diseases can be identified, as these are hypothesized to cluster in regions of low recombination. Such SNPs can serve as time-independent biomarkers of pathway involvement that may mechanistically explain the established associations. For example, we were interested in shared mechanisms between adult-attained height and postmenopausal breast cancer and colorectal cancer risk, because height is a risk factor for these cancers. Using the SNP selection approach, we identified clusters of GWAS SNPs that were associated with adult-attained height and the risk of postmenopausal breast cancer and/or colorectal cancer. This systematic approach identified a limited number of clustered SNPs, which pinpoint potential shared mechanisms (*i.e.* Indian Hedgehog signaling) that may link together the complex phenotypes height, postmenopausal breast cancer risk and/or colorectal cancer risk.

# 2. Adult-attained height and postmenopausal breast cancer: methodological considerations

In studying the height-postmenopausal breast cancer association, some specific methodological considerations are of interest, specifically the quality of the height measurements and to what extent height covers aspects of growth.

## 2.1 Quality of the height measurements

In the Netherlands Cohort Study, a large-scale prospective study with participants aged 55 to 69 years at baseline, height was self-reported at baseline. Adult-attained height is, in contrast to other anthropometric measurements, generally a stable measure over time. In addition, height is easy to acquire via self-reports. It is well-established that study participants are able to estimate their adult-attained height as strong correlations (>0.9) have been described between self-reported and measured height.[1-3] However, it cannot be excluded that some non-differential misclassification occurred in our cohort or other studies using self-reported height measurements. This misclassification can be either systematic as women tend to overestimate their height,[1,2] which could have attenuated the association between adult-attained height and postmenopausal breast cancer risk, or there could be regression to the mean, especially in the extremes of the height distribution, also contributing to attenuated associations.

## 2.2. Growth and development in early life

Adult-attained height is determined following a chronological order of events. Longitudinal growth starts in prenatal life and continues into the next two decades of life, characterized by periods of rapid growth, *i.e.* the growth spurt in infancy, the less intense mid-growth spurt in childhood, which is subject of debate in the literature,[4-6] and the pubertal growth spurt. The periods in between those growth spurts are generally marked by periods of relatively steady childhood growth. To give an indication of the relative contribution of growth phases over time to adult-attained height: a healthy individual reaches approximately 90% of his or her adult-attained height at the age at peak height velocity in puberty.[7] Since we discussed adult-attained height in relation to postmenopausal breast cancer risk, we continue focusing on longitudinal growth in girls. Roughly two years after the onset of the pubertal growth spurt, and just before the onset of menarche, the period of the highest growth rate during puberty is reached, known as age the peak height velocity.[8] Hereafter, the rate of growth is steadily reduced over time

until the maximum height is attained. The age at peak height velocity (as well as the peak height velocity itself), menarche, and maximum height attained may vary between individuals. In **Figure 1**, a schematic representation is given with regard to the course of longitudinal growth over time, in which the increase in height (**left box**) is presented relative to the corresponding height velocity (**right box**).

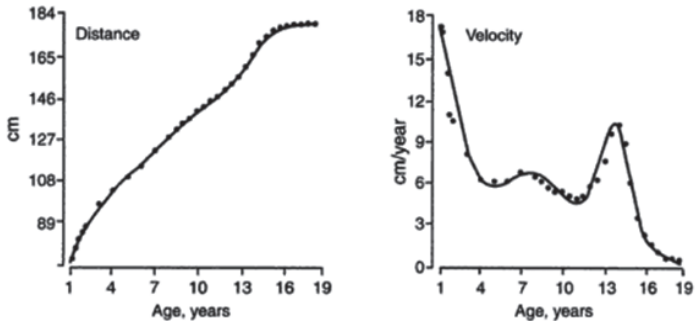



**Figure 1. A schematic representation of a mathematically fitted growth curve for stature of an individual child.** The circles indicate the measured statures (left) and respective height velocities (right). While we do not intend to focus on boys, this figure schematically depicts longitudinal growth in a boy. *Reprinted with permission from Taylor & Francis Online, source: M. el Lozy, 1978, ''A critical analysis of the double and triple logistic growth curves." Annals of Human Biology 5:389-394.*

Linear growth does not occur in a uniform manner throughout the body (**Figure 2**). *Tanner et al.,* observed a distal-to-proximal growth gradient, with more growth in distal body parts earlier in life than proximal body parts.[4,5,9-11] In this respect, the relative increase in leg length is largest before puberty and earlier in puberty, while the relative increase in trunk length is largest later in puberty and post puberty.[12,13] In women, the increasing estrogen levels in the body offset the ossification of the growth disks and this process will finally result in the closure of the growth disks, thereby marking the end of linear growth of the long bones, which include the legs.[14,15] Within approximately one year after the onset of the menarche, the growth disks are closed, marking the end of the adolescent growth spurt.

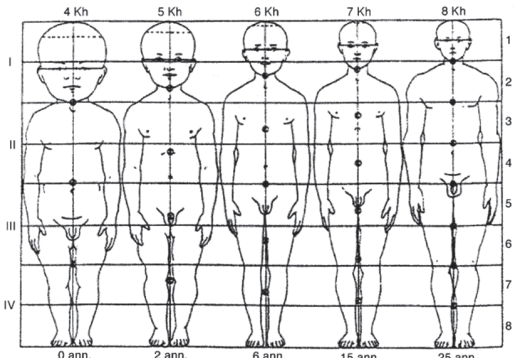



**Figure 2. Change in shape of the human body from birth to adulthood.** Between birth and puberty, the legs grow relatively faster than the upper body.[11,13] *Reprinted with permission from Oxford University Press, source: Prog Fd Nutr Sci (2) Leitch I. Change in shape of the human body, pp.99–141.*

## 2.3 Growth and development in relation to adult-attained height and postmenopausal breast cancer risk

### 2.3.1 Leg length and trunk length

In the Netherlands Cohort Study, data on adult-attained height were available, but no data on leg and trunk length of participants. Adult-attained height is correlated with both leg length ($r = 0.84$, $P < 0.001$) and trunk length ($r = 0.72$, $P < 0.001$), as for example shown in the EPIC study.[16] By studying the literature on leg length and trunk length in relation to postmenopausal breast cancer risk, we may gain further insight into the mechanisms underlying the height-cancer association. Although several cohort studies have observed that a larger leg length rather than a larger trunk length is associated with an increased postmenopausal breast cancer risk,[17-19] the EPIC study and other cohort studies found both leg length and trunk length to be associated with an increased postmenopausal breast cancer risk.[16,20,21] Only two of these studies looked into hormonal receptor subtypes of postmenopausal breast cancer. A longer leg length may particularly be associated with an increased risk of estrogen receptor-positive breast cancer.[18] The EPIC study, examining joint estrogen and progesterone receptor status, found that both a longer leg length and longer trunk length were associated with an increased risk of estrogen and progesterone receptor-positive breast cancer in women aged >60 years.[16]

### 2.3.2 Age at peak height velocity, age at menarche and age at maximum attained height

The age at peak height velocity,[22,23] age at menarche[24,25] and, age at maximum height[9,26,27] attained have been investigated in relation to adult-attained height and postmenopausal breast cancer risk. These associations may also provide further insight into the mechanisms underlying the height-cancer association.

Age at peak height velocity, age at menarche, and age at maximum attained height seem interrelated as women with an earlier menarche generally have an earlier age at peak height velocity[22] and an earlier age at maximum attained height.[28] Studies show that an earlier age at peak height velocity,[22] an earlier age at menarche,[25,29,30] and an earlier age at maximum attained height[25,28,30,31] may all be related to a shorter adult-attained height. The observed associations are, however, rather weak with the exception of the association between age at menarche and adult-attained height in the EPIC study,[25] which was quite strong. Biologically, this relationship seems plausible as an earlier menarche results in an earlier closure of the epiphyseal growth disks, thereby reducing the period of longitudinal growth of the long bones, including the legs, which may result in a shorter adult-attained height.[25] Nevertheless, there are also studies showing that an individual's adult-attained height is not associated with the timing of the pubertal growth spurt and, similarly, the age at menarche, age at peak height velocity, and age at maximum attained height in women.[6,32,33] In these studies, a lower peak height velocity (not to confuse with a younger age at peak height velocity) and smaller height gain was observed in girls with a later onset of the pubertal growth spurt compared to those with an earlier onset, which resulted in a similar adult-attained height in early and late maturing girls.[6,32,33] In the Netherlands Cohort Study, age at menarche was also not significantly correlated with adult-attained height (data not shown).

Findings indicate that an earlier age at peak height velocity may be associated with an increased postmenopausal breast cancer risk.[22] No studies are available on this association by hormone receptor status. An early age at menarche has also been related to an increased risk of postmenopausal breast cancer and this association is well established.[24] In addition, a meta-analysis showed that a later age at menarche was significantly associated with a decreased risk of both hormone receptor-positive and -negative breast cancer, yet the protective effect of a later age at menarche was statistically significantly stronger for hormone receptor-positive breast cancer.[34] In the Netherlands Cohort Study, age

at menarche was inversely associated with the risk of postmenopausal breast cancer overall, estrogen receptor-positive, progesterone receptor-positive, and estrogen receptor-negative breast cancer but not progesterone receptor-negative breast cancer (data not shown). The effect of an early age at menarche on the risk of hormone receptor subtypes of postmenopausal breast cancer has also been studied in interaction with adult-attained height. Women who were tall and had an early menarche (aged ≤13 years) showed an almost 2-fold increased risk of hormone receptor-positive tumors but no such increase in risk was observed for hormone receptor-negative tumors.[16]

With regard to an earlier age at maximum attained height, associations with an increased postmenopausal breast cancer risk have been reported.[26,27,35,36] Only one study investigated this association by hormone receptor status. This study showed that women who reached their maximum height at ≤12 years of age had an increased risk of estrogen receptor-negative breast cancer compared with women who reached their maximum height at ≥17 years of age.[26]

Taken together, it is noteworthy that the above variables, except for the age at maximum attained height, were particularly associated with hormone receptor-positive breast cancer. Although there are only a few studies to date, this may point to the involvement of hormonal (growth) factors in the observed associations. There is one notable contradiction in reviewing the literature between these exposures and breast cancer. Age at peak height velocity, age at menarche, and age at maximum attained height have all been associated with a shorter adult-attained height and an increased postmenopausal breast cancer risk, while tallness is also associated with an increased postmenopausal breast cancer risk. Height is more than a marker of these other variables in relation to breast cancer. As postulated in the EPIC study, the relative contribution of endogenous sex hormones and growth hormones to growth processes influencing postmenopausal breast cancer risk later in life may explain these seemingly contradictory findings.[25] This implies that adult-attained height also has an independent association with postmenopausal breast cancer risk.

# 3. Using early life exposures to shed light on the height-cancer association

It has been hypothesized that early life energy restriction may reduce adult-attained height. In addition, early life energy restriction has been known to reduce cancer risk. It is possible that energy restriction influences adult-attained height through (a) similar mechanism(s) that link adult-attained height to cancer risk.

## 3.1 Exposure assessment

Early life determinants of growth are presumably associated with cancer risk later in life in a direction as expected based on analogy with the height-cancer association. With regard to early life energy restriction, it should be mentioned that this is a unique exposure available within only a few cohorts worldwide.[37] In most cohort studies, exposure to early life energy restriction is war-related and proxy-assessed using information on residential status from self-reports or registries.[37] Self-reports were mostly based on questions relating to area-exposure data for a period of presumed energy restriction,[38-44] though one study has collected self-reports based on questions as to whether individuals can recall if they had experienced hunger during a period of presumed energy restriction.[45] The registries that were used contain information on migration status to approximate whether individuals experienced energy restriction or not.[46,47]

In the Netherlands Cohort Study, place of residence during the Hunger Winter, place of residence during the War Years, and employment status of an individual's father during the Economic Depression were used as proxies for early life energy restriction. It has been proposed that individual famine data may be more accurate than exposure data based on residential status.[48] In follow-up measurements of the female subcohort members of the Netherlands Cohort Study, participants were asked if they had really experienced hunger during the Hunger Winter. Of the women who reported severe hunger, 80% lived in a western city during this winter.[41] These results indicate that place of residence during the Hunger Winter is an adequate proxy for exposure to energy restriction. It has also been documented that severe energy restriction was confined to the western (famine) cities during the Hunger Winter.[49] Reports on this famine having effects on reproductive outcomes, birth weight, malformations, and perinatal mortality corroborate the severity of the energy restriction.[50] It has further been documented that food supplies deteriorated much faster in the cities than rural areas during the War Years and that a lower energy intake was associated with

an unemployed father during the Economic Depression.[50] These findings indicate that the proxies used are reasonably adequate and any misclassification is likely to be non-differential.

## 3.2 Energy restriction in relation to growth and development

It has been shown that energy restriction occurring early in life throughout growth and development may lead to decreased (hormonal) growth factor levels,[51] which in turn may result in a shorter adult-attained height.[52-54] In the Netherlands Cohort Study, only energy restriction during the Economic Depression resulted in a shorter adult-attained height (Chapter 3 of this thesis). Studying energy restriction in relation to leg length and trunk length may provide insight into the timing of exposures that influence growth and development and, thereby, adult-attained height and postmenopausal breast cancer risk. The rationale for such an approach lies in the observation that the pre- and peripubertal increase in linear growth is determined for a (much) larger part by an increase in leg length than trunk length,[12,19,55-57] and the observation that harsh conditions (in terms of energy restriction) at this time causes a shorter stature due to relatively short legs.[55-59] This is why leg length has been suggested to be a more sensitive measure of pre- and peripubertal exposures than trunk length.[12,19,55-57] By contrast, peak trunk length growth is generally reached after the age at peak height velocity,[9] and may, therefore, be a better measure of exposures later in puberty as well as (shortly) after puberty until an individual reaches his or her maximum height.[12,60]

The age at peak height velocity, age at menarche, and age at maximum attained height may also act as markers for exposures operating at different stages during growth.[56,61] In developing countries, undernutrition is associated with a later age at peak height velocity and age at menarche.[62] Other studies on energy restriction and age at menarche in Europe also showed delays in menarche after exposure to early life energy restriction.[25,54,63,64] In the Netherlands Cohort Study, early life energy restriction tended to significantly delay the age at menarche in women exposed to energy restriction before and or during the growth spurt (data not shown). Women who were severely exposed to energy restriction during the Hunger Winter (residing in a Western city), before and or during their growth spurt, had on average a later age at menarche (aged 15.6 years) compared to women exposed after the growth spurt (aged 13.3 years). Women who were exposed to energy restriction during the War Years (residing in an urban area in 1942), before and or during their growth spurt, had on average a later age at

menarche (aged 14.6 years) compared to women exposed after the growth spurt (aged 13.2 years). Due to the limited number of women exposed after the growth spurt it was not possible to determine this for energy restriction exposure to the Economic Depression.

## 3.3 Energy restriction and postmenopausal breast cancer risk

An apparent inconsistency in the findings discussed in this thesis is the discrepancy in the findings on energy restriction in the meta-analysis and those in the Netherlands Cohort Study. Early life energy restriction was associated with an increased postmenopausal breast cancer risk in the meta-analysis[37] as opposed to the generally protective effects observed in the Netherlands Cohort Study (Chapter 3 of this thesis).

One explanation for this discrepancy may be that we were only able to investigate the risk of postmenopausal breast cancer overall in the meta-analysis, while in the Netherlands Cohort Study we were able to perform the analyses according to hormone receptor subtypes. It is known that the distribution of estrogen and progesterone receptor subtypes among postmenopausal breast cancer cases may differ between study populations and that associations between risk factors and postmenopausal breast cancer may vary according to these subtypes. In the Netherlands Cohort Study, we observed that associations between early life energy restriction and postmenopausal breast cancer risk may indeed differ for hormone receptor-positive and -negative breast cancer. That we were not able to differentiate by hormone receptor subtype in the meta-analysis may thus have affected the strength and direction of the observed association.

Timing of exposure to early life energy restriction may be another factor that explains in part the discrepancies between the findings in both studies. In the meta-analysis, we were not able to disentangle the effect of timing of exposure from that of the duration and severity of exposure, while in the Netherlands Cohort Study we were able to investigate the influence of timing of exposure to energy restriction on postmenopausal breast cancer risk. Results showed that early life energy restriction was only associated with a reduced risk of hormone receptor-positive breast cancer in women who were exposed before and/or during the growth spurt.

Residual confounding due to unmeasured factors such as socio-economic status,[46,57] war-related stress,[65,66] malnutrition[67,68] and, comorbidities[50,69-71] may also be partly responsible for the observed increased postmenopausal breast cancer risk in our meta-analysis. Certainly, as most studies in the meta-analysis investigated severe famine and war-related energy restriction, other war-related factors are also likely to play a role, *i.e.* stress, malnutrition, and other famine- and war-related comorbidities. Some of these residual confounding factors may inflate a positive association. For instance, while it has been widely established that early life energy restriction reduces adult-attained height,[53,54] a study examining adult-attained height in women who survived the Siege of Leningrad indicated that these women were probably on average taller than those who did not survive the Siege.[46] It is possible that women who survived the Siege were part of a select group of women with a generally higher socio-economic status. A higher socio-economic status in turn is linked with a taller stature, probably because these women had more opportunities to increase their rationing during the Siege and as a consequence were able to survive during this period of adverse environmental circumstances.[46,70] In the Netherlands Cohort study, exposure to energy restriction during the Hunger Winter probably coincided with exposure to similar confounding factors. Although we observed a protective effect on hormone receptor-positive breast cancer in those women who were exposed before and/or during puberty, we also observed an increased risk of estrogen receptor-negative breast cancer, which was independent of timing of exposure. This may point to the involvement of other war-related factors. This is highly speculative though since the other proxies of early life energy restriction were not associated with increased risks for estrogen receptor-negative breast cancer.

# 4. Selecting genetic variants that link height to cancer risk

## 4.1 Involvement of hormonal mechanisms

In this thesis, a number of developmental growth variables have been discussed and it appears that most of these factors, including adult-attained height and early life exposures, *i.e.* energy restriction, may particularly be associated with hormone receptor-positive subtypes of postmenopausal breast cancer risk. This suggests possible hormonal links between adult-attained height and postmenopausal breast cancer risk.

Some of the SNPs identified from the SNP selection strategy as described in Chapter 4 also point to the involvement of hormonal mechanisms linking adult attained height to postmenopausal breast cancer risk, despite the fact that the GWAS studies included in this study did not specifically distinguish between estrogen and/or progesterone receptor status of breast cancer. For instance, some of the identified SNPs were annotated to genes involved in androgen receptor signaling, *i.e.* the *Estrogen Receptor-1* (*ESR1*) gene and the *Zinc Finger MIZ-Type Containing 1* (*ZMIZ1) gene*. Androgen receptor signaling can contribute to the development and structure of the adult mammary gland by the regulation of homeostasis between estrogens and androgens, which is essential to constraining the proliferative effects of estrogens.[72] In addition, androgen receptor signaling is also biologically relevant with regard to adult-attained height through its role of directly regulating the estrogenic effects on epiphyseal growth and maturation.[73]

Of specific interest is that the SNP selection strategy resulted in only one cluster that contained SNPs associated with all three phenotypes of interest, *i.e.* adult-attained height, postmenopausal breast cancer risk and colorectal cancer risk. Those SNPs were annotated to the *ZMIZ1 gene*. In preliminary results from a manuscript in preparation, a cluster-specific risk score for the *ZMIZ1* gene on chromosome 10 (summing the risk alleles from the GWAS catalog) was correlated with adult-attained height and associated with an increased risk of both estrogen and progesterone receptor-positive postmenopausal breast cancer in the Netherlands Cohort Study, but not with estrogen and progesterone receptor-negative postmenopausal breast cancer. Other studies also reported that one of the SNPs (rs704010) in chromosome 10 was associated with increased risk of estrogen receptor-positive[74,75] and progesterone receptor-positive breast cancer subtypes.[74] In the preliminary results from the Netherlands Cohort Study, the cluster-specific risk score of chromosome 10 was also associated with an increased distal colon cancer risk, which was confined to women (unpublished data). This result is in line with the significant association previously observed between adult-attained height and distal colon cancer which was also confined to women in the Netherlands Cohort Study.[76]

## 4.2 Future recommendations for using GWAS to select genetic variants

The preliminary results from the cluster-specific risk score of chromosome 10 emphasize that associations between adult-attained height and cancer risk are possibly gender-specific, subtype specific (*e.g.* hormone receptor status of

breast cancer), and dependent on anatomical subsites of cancer. To date, few GWAS studies have distinguished between gender, cancer subtype or subsite. GWAS studies, with stratification on these factors, may reveal genetic variants from other relevant pathways linking height to subtype-specific postmenopausal breast cancer[77] and/or colorectal cancer at different subsites[78] and gender specific associations[79,80].

In Chapter 4, the focus was on adult attained height, but there is also potential genetic variation in other measures of growth (paragraph 2.2 of this discussion) which could provide complementary and additional insights in relation to cancer endpoints compared to those obtained by studying height alone. One aspect is that there are different components to height (*e.g.* trunk and leg length). Although the genetic variation associated with height has been studied extensively in GWAS,[81,82] little is known about the genetic variation associated with the individual components of height and their relative importance (*e.g.* sitting-to-height ratio). Only one GWAS investigated these aspects.[83] In this genome-wide analysis of sitting-to-height ratio, performed using data from the GIANT consortium, a pathway analysis of height-related variants associated with the sitting-to-height ratio, and especially those associated with leg length, pointed to other biological pathways (*e.g.*, bone/cartilage/growth plate pathways), than when doing this analysis with height-related variants not associated with the sitting-to-height ratio (*e.g.*, embryonic development).[83]

Studies have also identified genetic variation relating to pre- and pubertal growth,[84-86] timing of the growth spurt,[84,85,87,88] peak height growth velocity,[89-91] and age at menarche,[84,85,92-94] with the exception of age at maximum attained height. Genetic association analyses in relation to different growth variables may be useful as these markers of growth and development may capture in part different pathways that regulate growth. This is illustrated by the observation that different variants were associated with the timing of the pubertal growth spurt and the timing of peak height velocity in infancy and in puberty in a prospective cohort.[91] Interestingly, this study also tested these variants for interaction with age (infancy versus puberty) and found evidence for an age-dependent effect for two SNPs. For instance, one of these SNPs had an effect on the peak height velocity in infancy but not on the peak height velocity in puberty. These observations may lend support to the concept of differences in the genetic regulation of human height during different periods of growth.

### 4.3 Identifying genetic variants as markers of pathway involvement

In this thesis, we identified genetic variants that can serve as markers for the pathways through which adult-attained height is linked with cancer risk, and accounting for these pathways may reveal associations that remain absent in overall analyses. Although such analyses are planned for the future with the selected SNPs in chapter 4, in a previous study from the Netherlands Cohort Study, adult-attained height was associated with a higher colorectal cancer risk in men only in the presence of an accumulation of unfavorable alleles in IGF-related genes.[95] This finding additionally illustrates how genetic variants can be used as markers of pathway involvement, *e.g.* the IGF-pathway in linking height to colorectal cancer in men. Our approach was aimed at identifying relevant pathways linking exposure to outcome, should not be confused with a Mendelian randomization approach. A Mendelian randomization approach aims to determine causal relationships between exposures and outcomes and is analogous to a randomized controlled trial.[96-99] Importantly, the instrumental variables used to obtain causal estimates of associations between an exposure, *i.e.* the SNPs associated with the exposure, should not be directly associated with the outcome.[100,101] Corroborating observational data, Mendelian randomization studies have shown associations between height and both postmenopausal breast cancer and colorectal cancer risk.[102,103]

## 5. Concluding remarks

The findings reported in this thesis require confirmation in other studies, but our observations support the growing body of evidence that the factors that contribute to a greater longitudinal growth, reflected by a greater adult-attained height, are associated with postmenopausal breast cancer risk and particularly estrogen and progesterone receptor-positive subtypes. In analogy with these findings, early life energy restriction before and or during the pubertal growth spurt both reduce adult-attained height and decrease the risk of estrogen and progesterone receptor-positive subtypes of postmenopausal breast cancer. Furthermore, the identified pathways using genetic variation derived from the SNP selection strategy supported the involvement of, among others, hormonal mechanisms linking adult attained height to both postmenopausal breast cancer risk and colorectal cancer risk. Therefore, the findings from this thesis point to: a) possible common underlying mechanism that involves hormonal growth factors that may link adult-attained height to postmenopausal breast and colorectal cancer risk;

and b) critical exposure periods in early life (before and during the pubertal growth spurt) when exposures have their prime effect on postmenopausal breast cancer risk. Therefore, cancer prevention strategies should take into consideration disease heterogeneity as indicated by molecular and hormonal subtypes and the anatomical location of the tumour. Specific tumour subtypes and tumors occurring in different anatomical subsites are often characterized by different risk factor profiles. In addition, the effects on adult-attained height and cancer risk may be lasting, even though early life exposures may be transient in nature, such as energy restriction. Therefore, studying early life modifiable exposures that influence adult-attained height and taking into account their timing in early life, could provide us with a better insight into when cancer prevention strategies are most effective. This may lead to new future cancer prevention strategies aimed at early life in contrast to the conventional prevention strategies that are generally focused on adults.

# References


1 Wright, F. L. *et al.* Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. *BMC Med Res Methodol* **15**, 81, doi:10.1186/s12874-015-0075-1 (2015).

2 Spencer, E. A., Appleby, P. N., Davey, G. K. & Key, T. J. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr* **5**, 561-565, doi:10.1079/PHN2001322 (2002).

3 Kuczmarski, M. F., Kuczmarski, R. J. & Najjar, M. Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. *J Am Diet Assoc* **101**, 28-34; quiz 35-26, doi:10.1016/S0002-8223(01)00008-6 (2001).

4 Tanner JM. Growth at adolescence. 2. Oxford: Blackwell; 1962.

5 Tanner, J. M. & Cox, L. A. London Archive of Longitudinal Growth Data. *Ann Hum Biol* **13**, 178-182 (1986).

6 Tanner, J. M., Whitehouse, R. H., Marubini, E. & Resele, L. F. The adolescent growth spurt of boys and girls of the Harpenden growth study. *Ann Hum Biol* **3**, 109-126 (1976).

7 Sanders, J. O. *et al.* The Uniform Pattern of Growth and Skeletal Maturation during the Human Adolescent Growth Spurt. *Sci Rep* **7**, 16705, doi:10.1038/s41598-017-16996-w (2017).

8 Robert M. Malina, C. B., Oded Bar-Or. Growth, Maturation, and Physical Activity. Chapter 3 Somatic growth., 62.

9 Rao, S., Joshi, S. & Kanade, A. Growth in some physical dimensions in relation to adolescent growth spurt among rural Indian children. *Annals of Human Biology* **27**, 127-138, doi:Doi 10.1080/030144600282244 (2000).

10 Bogin, B., Smith, P., Orden, A. B., Varela Silva, M. I. & Loucky, J. Rapid change in height and body proportions of Maya American children. *Am J Hum Biol* **14**, 753-761, doi:10.1002/ajhb.10092 (2002).

11 Gillman, M. W., Gluckman, P. D. & Rosenfeld, R. G. Recent advances in growth research: nutritional, molecular and endocrine perspectives. Preface. *Nestle Nutr Inst Workshop Ser* **71**, IX-X, doi:10.1159/000345196 (2013).

12 Gunnell, D. *et al.* Height, leg length, and cancer risk: a systematic review. *Epidemiol Rev* **23**, 313-342 (2001).

13 Bogin, B. & Varela-Silva, M. I. Leg length, body proportion, and health: a review with a note on beauty. *Int J Environ Res Public Health* **7**, 1047-1075, doi:10.3390/ijerph7031047 (2010).

14 Georgiadis, E., Mantzoros, C. S., Evagelopoulou, C. & Spentzos, D. Adult height and menarcheal age of young women in Greece. *Ann Hum Biol* **24**, 55-59 (1997).

15 Helm, P., Munster, K. & Schmidt, L. Recalled menarche in relation to infertility and adult weight and height. *Acta Obstet Gynecol Scand* **74**, 718-722 (1995).

16 Ritte, R. *et al.* Height, age at menarche and risk of hormone receptor-positive and -negative breast cancer: a cohort study. *Int J Cancer* **132**, 2619-2629, doi:10.1002/ijc.27913 (2013).

17 Albanes, D., Jones, D. Y., Schatzkin, A., Micozzi, M. S. & Taylor, P. R. Adult stature and risk of cancer. *Cancer Res* **48**, 1658-1662 (1988).

18 Fagherazzi, G., Vilier, A., Boutron-Ruault, M. C., Clavel-Chapelon, F. & Mesrine, S. Height, sitting height, and leg length in relation with breast cancer risk in the E3N cohort. *Cancer Epidemiol Biomarkers Prev* **21**, 1171-1175, doi:10.1158/1055-9965.EPI-12-0130 (2012).
19 Gunnell, D. J. *et al.* Childhood leg length and adult mortality: follow up of the Carnegie (Boyd Orr) Survey of Diet and Health in Pre-war Britain. *J Epidemiol Community Health* **52**, 142-152 (1998).
20 Mellemkjaer, L. *et al.* Leg length, sitting height and postmenopausal breast cancer risk. *Br J Cancer* **107**, 165-168, doi:10.1038/bjc.2012.244 (2012).
21 Lawlor, D. A., Okasha, M., Gunnell, D., Smith, G. D. & Ebrahim, S. Associations of adult measures of childhood growth with breast cancer: findings from the British Women's Heart and Health Study. *Br J Cancer* **89**, 81-87, doi:10.1038/sj.bjc.6600972 (2003).
22 Ahlgren, M., Melbye, M., Wohlfahrt, J. & Sorensen, T. I. Growth patterns and the risk of breast cancer in women. *N Engl J Med* **351**, 1619-1626, doi:10.1056/NEJMoa040576 (2004).
23 Berkey, C. S., Frazier, A. L., Gardner, J. D. & Colditz, G. A. Adolescence and breast carcinoma risk. *Cancer* **85**, 2400-2409 (1999).
24 Collaborative Group on Hormonal Factors in Breast, C. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* **13**, 1141-1151, doi:10.1016/S1470-2045(12)70425-4 (2012).
25 Onland-Moret, N. C. *et al.* Age at menarche in relation to adult height: the EPIC study. *Am J Epidemiol* **162**, 623-632, doi:10.1093/aje/kwi260 (2005).
26 Li, C. I., Littman, A. J. & White, E. Relationship between age maximum height is attained, age at menarche, and age at first full-term birth and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* **16**, 2144-2149, doi:10.1158/1055-9965.EPI-07-0242 (2007).
27 Li, C. I., Malone, K. E., White, E. & Daling, J. R. Age when maximum height is reached as a risk factor for breast cancer among young U.S. women. *Epidemiology* **8**, 559-565 (1997).
28 Baer, H. J. *et al.* Adult height, age at attained height, and incidence of breast cancer in premenopausal women. *International Journal of Cancer* **119**, 2231-2235, doi:10.1002/ijc.22096 (2006).
29 Shangold, M. M., Kelly, M., Berkeley, A. S., Freedman, K. S. & Groshen, S. Relationship between Menarcheal Age and Adult Height. *Southern Med J* **82**, 443-445 (1989).
30 Biro, F. M. *et al.* Impact of timing of pubertal maturation on growth in black and white female adolescents: The National Heart, Lung, and Blood Institute Growth and Health Study. *J Pediatr-Us* **138**, 636-643, doi:10.1067/mpd.2001.114476 (2001).
31 De Stavola, B. L. *et al.* Childhood growth and breast cancer. *Am J Epidemiol* **159**, 671-682 (2004).
32 Ferrandez, A. *et al.* Longitudinal Pubertal Growth According to Age at Pubertal Growth Spurt Onset: Data from a Spanish Study Including 458 Children (223 Boys and 235 Girls). *J Pediatr Endocr Met* **22**, 715-726 (2009).
33 Vizmanos, B., Marti-Henneberg, C., Cliville, R., Moreno, A. & Fernandez-Ballart, J. Age of pubertal onset affects the intensity and duration of pubertal growth peak but not final height. *Am J Hum Biol* **13**, 409-416, doi:10.1002/ajhb.1065 (2001).
34 Ma, H., Bernstein, L., Pike, M. C. & Ursin, G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res* **8**, R43, doi:10.1186/bcr1525 (2006).

35 Bertrand, K. A., Gerlovin, H., Bethea, T. N. & Palmer, J. R. Pubertal growth and adult height in relation to breast cancer risk in African American women. *Int J Cancer*, doi:10.1002/ijc.31019 (2017).
36 Bodicoat, D. H. *et al.* Timing of pubertal stages and breast cancer risk: the Breakthrough Generations Study. *Breast Cancer Res* **16**, R18, doi:10.1186/bcr3613 (2014).
37 Elands, R. J. *et al.* A Systematic Literature Review and Meta-Regression Analysis on Early-Life Energy Restriction and Cancer Risk in Humans. *Plos One* **11**, e0158003, doi:10.1371/journal.pone.0158003 (2016).
38 Robsahm, T. E. & Tretli, S. Breast cancer incidence in food- vs non-food-producing areas in Norway: possible beneficial effects of World War II. *Br J Cancer* **86**, 362-366, doi:10.1038/sj.bjc.6600084 (2002).
39 Hughes, L. A. *et al.* Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. *Int J Epidemiol* **39**, 1333-1344, doi:10.1093/ije/dyq062 (2010).
40 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Energy restriction early in life and colon carcinoma risk: results of The Netherlands Cohort Study after 7.3 years of follow-up. *Cancer* **97**, 46-55, doi:10.1002/cncr.11052 (2003).
41 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Diet in adolescence and the risk of breast cancer: results of the Netherlands Cohort Study. *Cancer Causes Control* **10**, 189-199 (1999).
42 Dirx, M. J., van den Brandt, P. A., Goldbohm, R. A. & Lumey, L. H. Energy restriction in childhood and adolescence and risk of prostate cancer: results from the Netherlands Cohort Study. *Am J Epidemiol* **154**, 530-537 (2001).
43 Heinen, M. M., Verhage, B. A., Goldbohm, R. A., Lumey, L. H. & van den Brandt, P. A. Physical activity, energy restriction, and the risk of pancreatic cancer: a prospective study in the Netherlands. *Am J Clin Nutr* **94**, 1314-1323, doi:10.3945/ajcn.110.007542 (2011).
44 Schouten, L. J., van Dijk, B. A., Lumey, L. H., Goldbohm, R. A. & van den Brandt, P. A. Energy restriction during childhood and early adulthood and ovarian cancer risk. *Plos One* **6**, e27960, doi:10.1371/journal.pone.0027960 (2011).
45 Elias, S. G., Peeters, P. H., Grobbee, D. E. & van Noord, P. A. Transient caloric restriction and cancer risk (The Netherlands). *Cancer Causes Control* **18**, 1-5, doi:10.1007/s10552-006-0080-0 (2007).
46 Koupil, I. *et al.* Cancer mortality in women and men who survived the siege of Leningrad (1941-1944). *International Journal of Cancer* **124**, 1416-1421, doi:Doi 10.1002/Ijc.24093 (2009).
47 Keinan-Boker, L., Vin-Raviv, N., Liphshitz, I., Linn, S. & Barchana, M. Cancer incidence in Israeli Jewish survivors of World War II. *J Natl Cancer Inst* **101**, 1489-1500, doi:10.1093/jnci/djp327 (2009).
48 Elias, S. G., Peeters, P. H., Grobbee, D. E. & van Noord, P. A. The 1944-1945 Dutch famine and subsequent overall cancer incidence. *Cancer Epidemiol Biomarkers Prev* **14**, 1981-1985, doi:10.1158/1055-9965.EPI-04-0839 (2005).
49 Stein, Z., Susser, M., Saenger, G. & Marolla, F. *Famine and human development, the Dutch Hunger winter of 1944-1945.*, (Oxford University Pres, 1975).
50 Lumey, L. H. & Van Poppel, F. W. The Dutch famine of 1944-45: mortality and morbidity in past and present generations. *Soc Hist Med* **7**, 229-246 (1994).

51 Elias, S. G. *et al.* Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. *Int J Cancer* **108**, 628-630, doi:10.1002/ijc.11584 (2004).
52 Dahlmann, N. & Petersen, K. Influences of environmental conditions during infancy on final body stature. *Pediatr Res* **11**, 695-700, doi:10.1203/00006450-197705000-00014 (1977).
53 Angell-Andersen, E. *et al.* The association between nutritional conditions during World War II and childhood anthropometric variables in the Nordic countries. *Ann Hum Biol* **31**, 342-355, doi:10.1080/03014460410001685304 (2004).
54 Brundtland, G. H., Liestol, K. & Walloe, L. Height, weight and menarcheal age of Oslo schoolchildren during the last 60 years. *Ann Hum Biol* **7**, 307-322 (1980).
55 Gunnel, D. Commentary: Can adult anthropometry be used as a 'biomarker' for prenatal and childhood exposures? *Int J Epidemiol* **31**, 390-394 (2002).
56 Wadsworth, M. E. J., Hardy, R. J., Paul, A. A., Marshall, S. F. & Cole, T. J. Leg and trunk length at 43 years in relation to childhood health, diet and family circumstances; evidence from the 1946 national birth cohort. *Int J Epidemiol* **31**, 383-390, doi:DOI 10.1093/ije/31.2.383 (2002).
57 Gunnell, D. J., Smith, G. D., Frankel, S. J., Kemp, M. & Peters, T. J. Socio-economic and dietary influences on leg length and trunk length in childhood: a reanalysis of the Carnegie (Boyd Orr) survey of diet and health in prewar Britain (1937-39). *Paediatr Perinat Epidemiol* **12 Suppl 1**, 96-113 (1998).
58 Cole, T. J. The secular trend in human physical growth: a biological view. *Econ Hum Biol* **1**, 161-168, doi:10.1016/S1570-677X(02)00033-3 (2003).
59 Hauspie, R. C., Vercauteren, M. & Susanne, C. Secular changes in growth. *Horm Res* **45**, 8-17 (1996).
60 Whitley, E., Martin, R. M., Smith, G. D., Holly, J. M. & Gunnell, D. Childhood stature and adult cancer risk: the Boyd Orr cohort. *Cancer Causes Control* **20**, 243-251, doi:10.1007/s10552-008-9239-1 (2009).
61 Gunnell, D. Commentary: Early insights into height, leg length, proportionate growth and health. *Int J Epidemiol* **30**, 221-222 (2001).
62 Malina RM, Bouchard C, Bar-Or O. Growth, maturation, and physical activity, 2nd ed. Champaign, Ill, USA: Human Kinetics, 2004.
63 van Noord, P. A. & Kaaks, R. The effect of wartime conditions and the 1944-45 'Dutch famine' on recalled menarcheal age in participants of the DOM breast cancer screening project. *Ann Hum Biol* **18**, 57-70 (1991).
64 Fentiman, I. S., Allen, D. S. & Ellison, G. T. The impact of the Occupation of Guernsey 1940-1945 on breast cancer risk factors and incidence. *Int J Clin Pract* **61**, 937-943, doi:10.1111/j.1742-1241.2007.01288.x (2007).
65 Vin-Raviv, N., Dekel, R., Barchana, M., Linn, S. & Keinan-Boker, L. World War II-related post-traumatic stress disorder and breast cancer risk among Israeli women: a case-control study. *Int Psychogeriatr* **26**, 499-508, doi:10.1017/S1041610213002081 (2014).
66 Sharon, A., Levav, I., Brodsky, J., Shemesh, A. A. & Kohn, R. Psychiatric disorders and other health dimensions among Holocaust survivors 6 decades later. *Br J Psychiatry* **195**, 331-335, doi:10.1192/bjp.bp.108.058784 (2009).
67 Shasha, S. M. [Morbidity in the ghettos during the Holocaust]. *Harefuah* **141**, 364-368, 409, 408 (2002).

68 Hercshlag-Elkayam, O., Even, L. & Shasha, S. M. [Clinical manifestations of "Hunger Disease" among children in the ghettos during the Holocaust]. *Harefuah* **142**, 345-349 (2003).
69 Stanner, S. A. *et al.* Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study. *BMJ* **315**, 1342-1348 (1997).
70 Koupil, I. *et al.* Blood pressure, hypertension and mortality from circulatory disease in men and women who survived the siege of Leningrad. *Eur J Epidemiol* **22**, 223-234, doi:10.1007/s10654-007-9113-6 (2007).

71 Huang, C., Li, Z., Wang, M. & Martore, R. Early Life Exposure to the 1959-1961 Chinese Famine Has Long-Term Health Consequences. *Journal of Nutrition* **140**, 1874-1878, doi:10.3945/jn.110.121293 (2010).
72 Peters, A. A., Ingman, W. V., Tilley, W. D. & Butler, L. M. Differential effects of exogenous androgen and an androgen receptor antagonist in the peri- and postpubertal murine mammary gland. *Endocrinology* **152**, 3728-3737, doi:10.1210/en.2011-1133 (2011).
73 Mauras, N., Rogol, A. D., Haymond, M. W. & Veldhuis, J. D. Sex steroids, growth hormone, insulin-like growth factor-1: neuroendocrine and metabolic regulation in puberty. *Horm Res* **45**, 74-80 (1996).
74 Li, X. *et al.* Association of multiple genetic variants with breast cancer susceptibility in the Han Chinese population. *Oncotarget* **7**, 85483-85491, doi:10.18632/oncotarget.13402 (2016).
75 Han, M. R. *et al.* Evaluating 17 breast cancer susceptibility loci in the Nashville breast health study. *Breast Cancer* **22**, 544-551, doi:10.1007/s12282-014-0518-2 (2015).
76 Hughes, L. A. *et al.* Body size and colorectal cancer risk after 16.3 years of follow-up: an analysis from the Netherlands Cohort Study. *Am J Epidemiol* **174**, 1127-1139, doi:10.1093/aje/kwr247 (2011).
77 Kristensen, V. N. & Borresen-Dale, A. L. SNPs associated with molecular subtypes of breast cancer: On the usefulness of stratified Genome-wide Association Studies (GWAS) in the identification of novel susceptibility loci. *Mol Oncol* **2**, 12-15, doi:10.1016/j.molonc.2008.02.003 (2008).
78 He, J. *et al.* Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. *Cancer Epidemiol Biomarkers Prev* **20**, 70-81, doi:10.1158/1055-9965.EPI-10-0892 (2011).
79 Winkler, T. W. & Consortium, G. How to Detect Genetic Loci in the Presence of Heterogeneity between Men and Women? Methodological Issues and Results from Gender-specific Meta-analyses for Human Anthropometric Measures: The GIANT Consortium. *Genet Epidemiol* **34**, 955-955 (2010).
80 Heid, I. M. *et al.* To Stratify Or Not To Stratify: What Can Be Learned From Power Considerations And A Practical Genome-Wide Search On Sex-Difference In The GIANT Consortium. *Genet Epidemiol* **36**, 147-148 (2012).
81 Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-838, doi:10.1038/nature09410 (2010).
82 Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173-1186, doi:10.1038/ng.3097 (2014).

83 Chan, Y. *et al.* Genome-wide Analysis of Body Proportion Classifies Height-Associated Variants by Mechanism of Action and Implicates Genes Important for Skeletal Development. *Am J Hum Genet* **96**, 695-708, doi:10.1016/j.ajhg.2015.02.018 (2015).

84 Cousminer, D. L. *et al.* Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. *Hum Mol Genet* **22**, 2735-2747, doi:10.1093/hmg/ddt104 (2013).

85 Cousminer, D. L. *et al.* Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. *Hum Mol Genet* **23**, 4452-4464, doi:10.1093/hmg/ddu150 (2014).

86 Vrieze, S. I. *et al.* An assessment of the individual and collective effects of variants on height using twins and a developmentally informative study design. *PLoS Genet* **7**, e1002413, doi:10.1371/journal.pgen.1002413 (2011).

87 Widen, E. *et al.* Distinct variants at LIN28B influence growth in height from birth to adulthood. *Am J Hum Genet* **86**, 773-782, doi:10.1016/j.ajhg.2010.03.010 (2010).

88 Willemsen, R. H. & Dunger, D. B. Normal Variation in Pubertal Timing: Genetic Determinants in Relation to Growth and Adiposity. *Endocr Dev* **29**, 17-35, doi:10.1159/000438957 (2016).

89 Buyken, A. E., Karaolis-Danckert, N. & Remer, T. Association of prepubertal body composition in healthy girls and boys with the timing of early and late pubertal markers. *Am J Clin Nutr* **89**, 221-230, doi:10.3945/ajcn.2008.26733 (2009).

90 Silventoinen, K., Haukka, J., Dunkel, L., Tynelius, P. & Rasmussen, F. Genetics of pubertal timing and its associations with relative weight in childhood and adult height: the Swedish Young Male Twins Study. *Pediatrics* **121**, e885-891, doi:10.1542/peds.2007-1615 (2008).

91 Sovio, U. *et al.* Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966. *PLoS Genet* **5**, e1000409, doi:10.1371/journal.pgen.1000409 (2009).

92 Elks, C. E. *et al.* Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet* **42**, 1077-1085, doi:10.1038/ng.714 (2010).

93 Fernandez-Rhodes, L. *et al.* Association of adiposity genetic variants with menarche timing in 92,105 women of European descent. *Am J Epidemiol* **178**, 451-460, doi:10.1093/aje/kws473 (2013).

94 Tu, W. Z. *et al.* Associations Between Menarche-Related Genetic Variants and Pubertal Growth in Male and Female Adolescents. *J Adolescent Health* **56**, 66-72, doi:10.1016/j.jadohealth.2014.07.020 (2015).

95 Simons, C. C. *et al.* Body size, physical activity, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk. *Carcinogenesis*, doi:10.1093/carcin/bgv077 (2015).

96 VanderWeele, T. J., Tchetgen Tchetgen, E. J., Cornelis, M. & Kraft, P. Methodological challenges in mendelian randomization. *Epidemiology* **25**, 427-435, doi:10.1097/EDE.0000000000000081 (2014).

97 Evans, D. M. & Smith, G. D. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. *Annu Rev Genom Hum G* **16**, 327-350, doi:10.1146/annurev-genom-090314-050016 (2015).

98 Zheng, J. *et al.* Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep* **4**, 330-345, doi:10.1007/s40471-017-0128-6 (2017).

99 Boef, A. G. C., Dekkers, O. M. & le Cessie, S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol* **44**, 496-511, doi:10.1093/ije/dyv071 (2015).
100 Burgess, S., Daniel, R. M., Butterworth, A. S., Thompson, S. G. & Consortium, E. P.-I. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol* **44**, 484-495, doi:10.1093/ije/dyu176 (2015).
101 Burgess, S. & Thompson, S. G. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* **42**, 1134-1144, doi:10.1093/ije/dyt093 (2013).
102 Zhang, B. *et al.* Height and Breast Cancer Risk: Evidence From Prospective Studies and Mendelian Randomization. *J Natl Cancer Inst* **107**, doi:10.1093/jnci/djv219 (2015).
103 Khankari, N. K. *et al.* Association between Adult Height and Risk of Colorectal, Lung, and Prostate Cancer: Results from Meta-analyses of Prospective Studies and Mendelian Randomization Analyses. *Plos Med* **13**, doi:ARTN e1002118. 10.1371/journal.pmed.1002118 (2016).

# English summary

Our main objective as introduced in **Chapter 1** was to gain a deeper understanding of how height is associated with cancer risk later in life. Increased adult-attained height has been consistently associated with increased cancer risk in the literature, yet plausible mechanisms remain to be further elucidated. Although there is a broad understanding of how early life environmental and genetic processes contribute to adult-attained height, there is still little evidence on how these factors might link to cancer risk. As it has been observed that height increases the risk of a number of different types of cancer and because relative risk estimates in relation to height are very similar across different cancers and in different populations, a common mechanism might be at play. From an environmental perspective we aimed to study these underlying mechanisms by investigating how early life energy restriction during childhood and adolescence influences both adult-attained height and cancer risk. Early life determinants of growth are presumably associated with cancer risk later in life in a direction as expected based on analogy with the height-cancer association.

We also aimed to specifically study height in relation to the risk of postmenopausal breast cancer by hormone receptor subtypes. Breast cancer includes hormone-sensitive tumors, which may be susceptible to hormonal growth factors influencing adult-attained height and cancer risk. Finding an association between height and hormone receptor-positive breast cancer indicates involvement of hormonal growth factors as plausible mechanisms.

Furthermore, a SNP selection approach was applied to identify genetic variants and genes that may link height to cancer risk. We focused on postmenopausal breast cancer and colorectal cancer. Both types of cancer share a subset of risk factors and height has been identified as a convincing risk factor for these cancers according to the World Cancer Research Fund. Germline genetic variants are useful as markers of shared mechanisms between adult-attained height because these are time-independent markers of pathway involvement.

**Chapter 2** consists of a systematic literature review and meta-analysis of human observational studies in which the evidence on the association between early life energy restriction and site-specific cancer risk is investigated. Furthermore, specific aspects of early life energy restriction are discussed, such as the timing, duration, and severity of exposure, which may determine whether exposure is

associated with an increased or decreased risk of cancer. The systematic review indicated that moderate continuous energy restriction, as studied in animal experimental and human ecological studies, was generally associated with a decreased site-specific cancer risk. The meta-analysis of human observational studies indicated that severe transient early life energy restriction was associated with a 28% increased breast cancer risk and a 16% increased prostate cancer risk, though some of the underlying studies showed null results. The evidence for an association between severe transient early life energy restriction and risk at other cancer sites, *i.e.* colorectal-, stomach-, pancreas-, ovarian-, and respiratory cancer was either limited or studies were too heterogeneous for pooling. A subsequent meta-regression analysis investigating the effect of the duration and severity of energy restriction on overall cancer risk indicated that duration, rather than severity might result in increased overall cancer risk in women and men, however, this result should be interpreted with caution. With regard to timing of exposure to early life energy restriction, no conclusions could be drawn with regard to breast cancer risk.

In **Chapter 3** associations of adult-attained height and early life energy restriction with postmenopausal breast cancer risk are studied according to estrogen and progesterone receptor status within the Netherlands Cohort Study on Diet and Cancer. The Netherlands Cohort Study is a prospective cohort study among 120,852 participants, of which 58,279 men and 62,573 women and has data available on adult-attained height and the rather unique exposure of early life energy restriction. Adult-attained height was significantly positively associated with postmenopausal breast cancer risk, in particular with hormone receptor-positive subtypes. Of the three exposures to energy restriction investigated, *i.e.* exposure to energy restriction during the Hunger Winter (the winter of 1944-45), the War Years (1940-44), and the Economic Depression (1932-40), only exposure to energy restriction during the Economic Depression was related to a shorter stature (an almost 2 cm reduction) in female subcohort members. Energy restriction during all three periods of exposure, provided that the exposure occurred before and/or during the growth spurt, was associated with a significantly decreased risk of hormone receptor-positive breast cancer subtypes. Interestingly, energy restriction during the Hunger Winter increased the estrogen receptor-negative breast cancer risk regardless of the timing of energy restriction. Taken together, the observation that both height and early life energy restriction taking place before and/or during the growth spurt were associated with reduced hormone receptor-positive breast cancer risk seems to suggest possible common underlying mechanisms and critical exposure periods in life.

In **Chapter 4**, an approach for the selection of genetic variants is presented, which was developed to select genetic variants to identify shared mechanisms linking height to postmenopausal breast- and colorectal cancer risk. The selection method for genetic variants is based on the assumption that genetic variants from GWAS associated with complex diseases or traits tend to co-segregate in regions of low recombination, harboring functionally linked gene clusters. This phenomenon allows for selecting a limited number of genetic variants. This selection approach for genetic variants is of particular interest for large-scale studies with exhaustive bio-samples, *i.e.* DNA from nails, in which a genome-wide approach is not feasible, and will reduce the costs of genotyping and the chance of false-positive findings. The novelty of this method is the comprehensive integration of publically available GWAS repositories, on the basis of which genetic variants associated with complex traits and diseases can be identified, as these are hypothesized to cluster in regions of low recombination. Genomic regions including genetic variants for height and cancer may indicate pathways underlying height-cancer associations. Such genetic variants can serve as time-independent biomarkers of pathway involvement that may mechanistically explain the established associations. Using our selection method for genetic variants, we identified clusters of genetic variants derived from GWAS data that were associated with adult-attained height and the risk of postmenopausal breast cancer and/or colorectal cancer. This systematic approach identified a limited number of clustered genetic variants, which pinpoint potential shared mechanisms (*i.e.* Indian Hedgehog signaling) that may link together the complex phenotypes height, postmenopausal breast cancer risk and/or colorectal cancer risk.

Finally, **Chapter 5** concludes the thesis with a discussion of the main findings as well as the challenges of and insights offered by (molecular) epidemiology on studying the height-cancer association. The findings reported in this thesis require confirmation in other studies, but our observations support the growing body of evidence that the factors that contribute to a greater longitudinal growth, reflected by a greater adult-attained height, are associated with postmenopausal breast cancer risk and particularly estrogen and progesterone receptor-positive subtypes. In analogy with these findings, early life energy restriction before and or during the pubertal growth spurt both reduce adult-attained height and decrease the risk of estrogen and progesterone receptor-positive subtypes of postmenopausal breast cancer. Furthermore, the identified pathways using genetic variation derived from the novel selection method for genetic variants supported the involvement of, among others, hormonal mechanisms linking adult-attained

height to both postmenopausal breast cancer risk and colorectal cancer risk. Therefore, the findings from this thesis point to: a) a possible common underlying mechanism that involves hormonal growth factors that may link adult-attained height to postmenopausal breast and colorectal cancer risk; and b) critical exposure periods in early life (before and during the pubertal growth spurt) when exposures may have their prime effect on postmenopausal breast cancer risk. As indicated by this thesis and the literature, specific tumour subtypes and tumours occurring in different anatomical subsites are often characterized by different risk factor profiles. In addition, the effects of early life exposures on adult-attained height and cancer risk may be lasting, even though early life exposures may be transient in nature, such as energy restriction. Therefore, studying early life modifiable exposures that influence adult-attained height and taking into account their timing in early life, could provide us with a better insight into when cancer prevention strategies are most effective. This may lead to new future cancer prevention strategies aimed at early life in contrast to the conventional prevention strategies that are generally focused on adults.

## Valorization addendum

Our main objective was to gain a deeper understanding of how adult-attained height is associated with cancer risk later in life. In analogy, we also investigated early-life energy restriction exposures during critical times of growth in relation to height and cancer risk, with cancer risk occurring much later in life. In the meta-analysis conducted in this thesis we observed that severe transient early life energy restriction was associated with a 28% increased breast cancer risk and a 16% increased prostate cancer risk, in the meta-analysis of human observational studies in this thesis. A systematic review conducted in this thesis indicated that moderate energy restriction, as studied in animal experimental and human ecological studies, may be associated with a decreased site-specific cancer risk. With regard to timing of exposure to early life energy restriction, no conclusions could be drawn with regard to breast cancer risk in our meta-analysis. A subsequent meta-regression analysis investigating the effect of the duration and severity of energy restriction on overall cancer risk indicated that duration, rather than severity might result in increased overall cancer risk in women and men, however, this result should be interpreted with caution.

The height-cancer risk association is consistent in the literature and the results presented in this thesis contribute to elucidating plausible mechanisms for this association. We also investigated adult-attained height in relation to the risk of postmenopausal breast cancer by estrogen and progesterone receptor subtypes. Adult-attained height and hormone-receptor positive tumors are thought to share causal mechanisms, specifically hormonal growth factors. In this thesis, adult-attained height and energy restriction before and/or during the growth spurt were both associated with the risk of estrogen- and progesterone receptor-positive subtypes of postmenopausal breast cancer, in the direction as expected, indicating critical exposure windows for hormonal growth-related mechanisms.

Finally, we developed a novel molecular epidemiological approach to identify the underlying mechanisms that may link height to cancer risk. This selection method for genetic variants from GWAS offers a systematic strategy for large-scale prospective cohort studies, interested in studying the underlying mechanisms between multiple phenotypes, such as a quantitative trait (*e.g.* height) and complex diseases (*e.g.* breast cancer risk and colorectal cancer risk). We used existing repositories of results from genome-wide association studies (GWAS) to identify adjacently located genetic variants that were associated with adult-

attained height and breast or colorectal cancer. The methodology in this approach relies on the assumption that genetic variants from GWAS found associated with complex diseases or traits are not randomly distributed across the genome, but tend to cluster in regions of low recombination [1]. With this approach, we were able to narrow down the large number of genetic variants from GWAS associated with these phenotypes to a limited set of genetic variants clustered to the same gene(s). The genetic variants in these clusters may collectively point to genes in pathways explaining the link between height and cancer risk. The novelty of this method is the comprehensive integration of publically available GWAS repositories, on the basis of which genetic variants associated with multiple associated complex traits and diseases can be identified. The past decade has seen a large number of wide-scale genetic studies in identifying genetic variants that may modify individuals' predisposition to common diseases.[2] The GWAS catalog[3] has now grown to contain tens of thousands of SNPs associated with hundreds of common diseases. However, the interpretation of these variants lags far behind.[2] Efforts such as the method introduced in this thesis, may allow using this data to systematically select genetic variants in large-scale epidemiological studies investigating disease aetiology. Existing strategies prioritise genetic variants related to exposure in cases and controls [4] or genetic variants related to the outcome [5], in which genetic variants are prioritised based on a genome-wide scan in the own study population. The selection approach for genetic variants presented in this study is of particular interest for application in large-scale studies with exhaustive bio-samples, *e.g.* DNA from nails, in which a genome-wide scan within the own study population is not feasible.

In this valorization addendum, the results of this thesis will be discussed in the light of, "value for society" and "value for science".

## Value for society

### Attention for timing of exposure

With regard to the results in this thesis, it is obvious that adult-attained height in itself cannot be the target of future interventions. Adult-attained height can be seen as a result of exposures to growth factors. It is a marker for genetic, environmental, including nutritional factors, and hormonal factors affecting growth during the period from preconception to completion of linear growth. Energy restriction is one of the environmental factors that may negatively influence energy balance and thereby may oppose the effects of exposure to growth factors. When energy restriction occurs early in life, it could therefore reduce cancer risk in

adulthood. Yet, it seems that a differentiation with regard to energy restriction occurring early in life and cancer risk later in life has to be made based on the timing of exposure to early life energy restriction, particularly during periods of rapid growth and development, such as during the pubertal growth spurt, and its subsequent effect on adult-attained height and cancer risk. These periods also coincide with the timing of breast tissue growth and maturation. If this theory is true, this may hold implications for lifelong prevention of cancer. When starting primary prevention of cancer early in life instead of in adulthood, the effectiveness of prevention might be larger than currently anticipated. Our study suggests that primary prevention of cancer is a lifelong goal that has to start early in life covering critical times of growth and development.

### The relevance of understanding biological mechanisms for prevention

Another aspect to consider when developing prevention strategies may be that risk factors for postmenopausal breast cancer differ by hormone receptor-positive and hormone receptor-negative subtypes. Generally, in contrast to hormone receptor-negative subtypes, hormone receptor-positive breast cancer tumours, *i.e.* estrogen- and progesterone hormone receptor-positive tumours, have been associated with modifiable (hormonal) growth-related risk factors, which can be a target for prevention at a young age. Given that approximately 80% of all breast cancer tumours in the population of the Netherlands is estrogen hormone receptor-positive, of which 65% is also progesterone receptor-positive,[6] breast cancer prevention strategies aimed at modifiable (hormonal) growth-related risk factors may be particularly promising in reducing the incidence of hormone receptor-positive breast cancer. Our results suggest that when estimating the potential for prevention it is advisable to take the molecular subtypes of tumours reflecting biological mechanisms into account. The knowledge and insight into potential biological mechanisms provided in this thesis may contribute to the justification of prevention policies aimed at modifiable early life exposures.

## Value for science

It is important to understand which aspects of growth relate to cancer risk as it might expand our knowledge about the pathways that lead to cancer development. Height is a reflection of absolute linear growth, and while it may be the outcome of the entire growth curve, it does not reflect the entire growth curve well when it comes to the timing, duration, and velocity of growth, hormonal changes influencing growth, and exposure to general growth factors. In addition,

height does not provide information on the contribution of leg length relative to trunk length. This information could give important clues given that a short height due to relatively short legs is generally associated with different early life exposures, particularly with exposures related to a social adverse environment, [7] compared to exposures associated with a short height due to a relatively short trunk length. Likewise, early life energy restriction is just one factor associated with the negative end of the spectrum of energy balance, which, as a whole, is associated with different factors among which can also be discerned nutritional status, physical activity, and body composition. Early life energy restriction was operationalized through very specific historical events in this thesis, which generally represented extreme circumstances. Nevertheless, studies focusing on the negative end of the energy balance spectrum are important, as most epidemiological studies and also prevention have focused on the positive end of this spectrum, *i.e.* childhood obesity, and cancer risk.[8-11]

More insight into factors contributing to growth that lead to cancer development by investigating the determinants of adult-attained height, *e.g.* by investigating the growth curve and components of height as well as which aspects of energy restriction link height to cancer risk, as described in Chapter 5 of this thesis.

Further clues with regards to the mechanisms underlying the height-cancer association may be found using the framework of the collaborative effort between WCRF and Bristol University for systematic reviews of mechanisms underpinning exposure-cancer associations.[12] This collaboration has led to the development of a novel two-stage framework on the mechanisms underpinning the association between exposure and cancer risk that provides: (stage I) an overview of mechanistic pathways wherein results from human (*i.e.*, epidemiologic), animal, and cell studies are integrated and (stage II) a systematic review of the literature for identifying biologically plausible mechanisms underlying these exposure-cancer associations. While there are well-established methodologies for systematic reviews of epidemiologic data, this framework offers guidelines for performing a systematic review of mechanistic evidence underlying well-known epidemiologic associations, which has not been proposed previously. Both the selection method for genetic variants from GWAS described in this thesis and the mechanistic framework from Bristol University mentioned above can be used to identify potential mechanisms for underpinning exposure-cancer associations with regard to timing, duration and growth velocity, and for hormonal and growth factors. With respect to adult-attained height, genome-wide associations and

longitudinal analyses have already linked genetic loci to pubertal height growth, pubertal timing and childhood adiposity.[13,14]

The research conducted within this thesis is part of a larger molecular epidemiological research line at the Department of Epidemiology and the GROW-School for Oncology and Developmental Biology at Maastricht University. This research line focusses on exposures related to energy balance, including body size and gene-environment interactions, and molecular characteristics of tumours. The aim is to further elucidating the contribution of energy balance-related risk factors to cancer risk, including energy balance-related risk factors in early life, and the pathways underlying these associations. Next to cancer cohorts such as the Netherlands Cohort Study, birth cohort studies with cancer follow-up into old age with detailed data on linear growth and genetic data could yield valuable insights, but these studies are scarce. Likewise, epidemiologic data regarding the mechanisms underlying an association between early-life energy restriction and human site-specific cancer risk are scarce because exposure to energy restriction is rarely available in observational studies and few studies are large enough to allow for subgroup analyses for molecular and histological specific endpoints. Birth cohort studies with cancer follow-up into old age and detailed data on growth and energy balance in early life (with attention to the full spectrum of energy balance) are promising, but such studies are not mature enough yet.

In the meantime, identifying potential mechanisms linking height to cancer should be pursued and these mechanisms could be identified and investigated with respect to the height-cancer risk association in ongoing prospective cohort studies.

To conclude, the results of this thesis point to critical exposure windows for hormonal growth-related mechanisms for postmenopausal breast cancer. Generally, early life determinants of growth, particularly during critical periods of growth, seem to be relevant to cancer that occurs after 55 years of age. Understanding how height is related to cancer is important for public health prevention strategies in childhood and for expanding our knowledge about the pathways to cancer development.

# References


1 Preuss, C., Riemenschneider, M., Wiedmann, D. & Stoll, M. Evolutionary dynamics of co-segregating gene clusters associated with complex diseases. *PLoS One* **7**, e36205, doi:10.1371/journal.pone.0036205 (2012).

2 Wijmenga, C. & Zhernakova, A. The importance of cohort studies in the post-GWAS era. *Nat Genet* **50**, 322-328, doi:10.1038/s41588-018-0066-3 (2018).

3 Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* **42**, D1001-1006, doi:10.1093/nar/gkt1229 (2014).

4 Murcray, C. E., Lewinger, J. P. & Gauderman, W. J. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol* **169**, 219-226, doi:10.1093/aje/kwn353 (2009).

5 Kooperberg, C. & Leblanc, M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. *Genet Epidemiol* **32**, 255-263, doi:10.1002/gepi.20300 (2008).

6 https://borstkanker.nl.

7 Gunnell, D. J., Smith, G. D., Frankel, S. J., Kemp, M. & Peters, T. J. Socio-economic and dietary influences on leg length and trunk length in childhood: a reanalysis of the Carnegie (Boyd Orr) survey of diet and health in prewar Britain (1937-39). *Paediatr Perinat Epidemiol* **12 Suppl 1**, 96-113 (1998).

8 Shawon, M. S. R., Eriksson, M. & Li, J. Body size in early life and risk of breast cancer. *Breast Cancer Res* **19**, 84, doi:10.1186/s13058-017-0875-9 (2017).

9 Lope, V. *et al.* Perinatal and childhood factors and risk of breast cancer subtypes in adulthood. *Cancer Epidemiol* **40**, 22-30, doi:10.1016/j.canep.2015.11.004 (2016).

10 Sangaramoorthy, M., Phipps, A. I., Horn-Ross, P. L., Koo, J. & John, E. M. Early-life factors and breast cancer risk in Hispanic women: the role of adolescent body size. *Cancer Epidemiol Biomarkers Prev* **20**, 2572-2582, doi:10.1158/1055-9965.EPI-11-0848 (2011).

11 Robinson, W. R., Poole, C. & Godley, P. A. Systematic review of prostate cancer's association with body size in childhood and young adulthood. *Cancer Causes Control* **19**, 793-803, doi:10.1007/s10552-008-9142-9 (2008).

12 Ertaylan, G. *et al.* A Comparative Study on the WCRF International/University of Bristol Methodology for Systematic Reviews of Mechanisms Underpinning Exposure-Cancer Associations. *Cancer Epidemiol Biomarkers Prev* **26**, 1583-1594, doi:10.1158/1055-9965.EPI-17-0230 (2017).

13 Cousminer, D. L. *et al.* Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. *Hum Mol Genet* **22**, 2735-2747, doi:10.1093/hmg/ddt104 (2013).

14 Sovio, U. *et al.* Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966. *PLoS Genet* **5**, e1000409, doi:10.1371/journal.pgen.1000409 (2009).

# Curriculum Vitae

Rachel Elands was born in 's-Hertogenbosch, the Netherlands on the 6th of August 1987. After completing secondary school at Jeroen Bosch College in 's-Hertogenbosch, the Netherlands, in 2004, she obtained her propaedeutics in Applied Sciences in 2005 at Fontys in Eindhoven, the Netherlands. In 2006 she started her Bachelor studies in Bio-Pharmaceutical Sciences in Leiden, the Netherlands. During the bachelor program, Rachel participated in a 6-month internship in medicinal chemistry at the department of Modern Drug Discovery, Leiden Amsterdam Centre for Drug Discovery in Leiden (LACDR), the Netherlands. This research internship aimed at examining the mechanism of activation of G protein-coupled receptors, among which is the Adenosine $A_{2B}$ receptor, also known as 'the caffeine receptor', the findings were presented in a co-authored paper.

After obtaining the Bachelor degree, Rachel enrolled in a master program with a two-year research specialization in Bio-Pharmaceutical Sciences where she participated in a 9-month internship at the department of Medical Pharmacology, LACDR in Leiden, the Netherlands. The internship aimed to investigate the role of a protein, doublecortin-like, in the mitochondrial regulation of energy homeostasis in neuroblastoma cell proliferation; the findings were presented in a co-authored paper. Through this internship, contacts with a researcher from the Betula project, a longitudinal study on aging and memory were established based at Stockholm University in Stockholm, Sweden. Here, Rachel participated in a 6-month research internship. This internship was aimed at understanding the networks in the brain that may provide insight into the relation between emotional attentional regulation and cognitive decline in older adults. During her studies, Rachel was a board member and writer for the faculty journal of Mathematics and Natural Sciences, Leiden University and employed at Rijksmuseum Boerhaave, a museum on the history of Medical and Natural Sciences in Leiden. After graduating from the Research Master in August 2012, from October 2012, Rachel worked as a junior researcher at the department of Environmental Health at the Karolinska Institute in Stockholm, Sweden. As a junior researcher she studied the mechanisms of 'caspases', which are involved in programmed cell-death. During this time she participated in a public outreach event to connect the public with the research performed at the institute.

In June 2013, she started her PhD at the department of Epidemiology within the GROW School for Oncology and Developmental Biology at Maastricht University in Maastricht, the Netherlands. The goal of this PhD project was to have a better understanding of the mechanisms underlying the link between height and cancer risk, focusing on post-menopausal breast and colorectal cancer risk. Adult-attained height and early life energy restriction during critical time windows of growth were investigated in relation to breast cancer risk in the Netherlands, currently inhabiting the tallest population in the world, by making use of the Netherlands Cohort Study. In addition, she conducted a systematic literature review and meta-regression analysis on early-life energy restriction and cancer risk in humans. Finally, she developed a systematic selection approach for genetic variants to identify mechanisms underlying the association between height and post-menopausal breast and colorectal cancer. In 2015, Rachel won the poster price during the GROW Science Day when presenting the SNP selection approach. During her PhD, Rachel also was also involved in teaching.

# List of publications, presentations and posters

## Articles

- Associations of adult-attained height and early life energy restriction with postmenopausal breast cancer risk according to estrogen and progesterone receptor status. *Under revision at the International Journal of Cancer.* **Elands RJJ**., Offermans NSM., Simons CCJM., Schouten LJ., Verhage BAJ., van den Brandt PA., Weijenberg MP.
- A systematic SNP selection approach to shed light on the mechanisms underlying disease aetiology: the example linking height and risk of post-menopausal breast and colorectal cancer. *Scientific Reports, 2017.* **Elands RJJ**., Riemenschneider M., Simons CCJM., Isaacs A., Schouten LJ., Verhage BAJ., Van Steen K., Godschalk RWL., van den Brandt PA., Stoll M., Weijenberg MP. A systematic literature review on early life energy restriction and cancer risk. *PLoS One, 2016.* **Elands RJJ**., Simons CCJM., Schouten LJ., Verhage BAJ., van Dongen M., van den Brandt PA., Weijenberg MP.
- Domains for activation and inactivation in G protein-coupled receptors-a mutational analysis of constitutive activity of the adenosine $A_{2B}$ **receptor.** *Biochem Pharmacol*. 2014 Aug 29. Peeters MC., Li Q., **Elands RJJ**, van Westen GJ., Lenselink EB., Müller CE., IJzerman AP.
- Silencing of doublecortin-like (DCL) results in decreased mitochondrial activity and delayed neuroblastoma tumor growth. *PLoS One*. 2013 Sep 26;8(9). Verissimo CS., **Elands RJJ**, Cheng S., Saaltink DJ., ter Horst JP., Alme MN., Pont C., van de Water B., Håvik B., Fitzsimons CP., Vreugdenhil E.

## Presentations and posters

- A systematic SNP selection approach for studying gene-environment interactions in cancer epidemiology. *Healthy Living: The European Congress of Epidemiology, 2015, Maastricht, the Netherlands.* **Elands RJJ**., Riemenschneider M., Simons CCJM., Schouten LJ., Verhage B., Van Steen K., Godschalk RWL., Brandt P., Stoll M., Weijenberg MP. (oral presentation)
- **Elands RJJ**., Simons CCJM., Schouten LJ., Verhage BAJ., van den Brandt PA., Weijenberg MP. Early life energy restriction and cancer risk in humans: a systematic literature review and meta-regression analysis. *GROW Science Day 2015, Maastricht, the Netherlands*. (poster)

- **Elands RJJ**., Riemenschneider M., Simons CCJM, Schouten LJ., Verhage BAJ., Van Steen K., Godschalk R., van den Brandt PA., Stoll M., Weijenberg MP. A systematic SNP selection strategy on common genetic variants for height and post-menopausal breast and colorectal cancer risk. *GROW Science Day 2014, Maastricht, the Netherlands*. (poster)
- **Elands RJJ**, Simons CC., Schouten LJ., Verhage BAJ., Dongen M., Brandt PA., Weijenberg MP. A systematic literature review on early life energy restriction and cancer risk. Diet and Cancer Symposium: from prevention to survival. *VLAG International Advanced Master Class, October 2nd 2014, Wageningen, the Netherlands*. (poster)
- **Elands RJJ**, Simons CC., Brandt PA., Schouten LJ., Verhage BAJ., Weijenberg MP. Early life energy restriction and cancer risk: a systematic literature review of observational studies in humans. *WEON: 38th National Epidemiology Conference, June 6th 2014, Leiden*. (poster)
- Ziaei M., **Elands RJJ**., Peira N., Persson J. Fusiform & Inferior Frontal Gyrus activation in attentional top-down modulation to emotional stimuli. *OHBM 18th Annual Meeting, June 10-14, 2012, Beijing, China*. (poster)

# Dankwoord

During the past four years, I have changed as a person. The training of my mind to critically assess my own views and thought processes, I owe to many people.

First of all, I would like to express my special thanks to my promotion team, who helped me to pave my research path. I am sincerely thankful to my promotor *Prof. Dr. Matty Weijenberg* for all the kind guidance and inspiring and stimulating discussions we had until the very end of my PhD journey. As my promotor it has probably not always been easy for you to guide me, as I am a bursting waterfall of ideas and I am keen on doing things in my own particular way. I appreciate your trust in me and the freedom to enter unconventional roads within my PhD project, starting with the fact that I was not trained as an epidemiologist. I would also like to thank my co-promotor *Dr. Colinda Simons* for your thorough review of my work throughout my PhD and our comprehensive discussion sessions. I would also like to thank my co-promotor *Prof. Dr. Piet van den Brandt* for your practical input, and allowing me to develop in a critical, collected researcher.

My thanks also goes to *Monika Stoll*, with her calm go-get outlook I basically needed just 10 minute-long conversations to figure out how to go ahead with developing the unconventional SNP selection strategy. I am grateful to *Mona Riemenschneider* for introducing me into the world where bioinformaticians meet epidemiologists and her help with improving the SNP selection method in every way. I also want to thank *Shirin Glander, Roger Godschalk, Kristel van Steen and Aaron Isaacs* for their detailed insights and feedback on this work. In addition, I would like to thank *Leo Schouten, Martien van Dongen* and *Nadine Offermans* for their critical review of my papers.

I also want to thank *Bas Verhage* for not only contributing to the research performed within this thesis but also for introducing me to the PBL system, which was totally different from the education systems I was acquainted with. Besides Bas, I would also like to thank *Huub Hoofs* for investing a lot of time in creating tools that aid tutors in making the sessions with the students more interactive and dynamic, I benefitted from such tools during my sessions as a tutor.

I would like to express my thanks to the help of all the administrators at the department of Epidemiology: *Yvonne Leenders, Yvonne van den Berg, Marielle van Os and Petra Brouns*, who did their best to smooth out bureaucratic matters,

which was indispensable because I also tend to order things in my own way. I would also like to thank *Harry van Montfort* and *Jos Slangen* for their ICT support.

Moreover, my colleagues in Maastricht have contributed immensely on my professional and personal life. I would like especially to thank *Eline van Roekel*, for all the heartfelt talks and listening to me even when you didn't have enough time for yourself. Furthermore, I would like to thank *IJmert Kant* who I got to know very well during a kayak activity on the annual day-out of the Epidemiology department. From that day on, he checked in at my office on most Friday evenings to tell me I had to leave the office but not before sharing me some of his latest mantra's of life. I want to thank him for his advice and inspirational words until the very end of my PhD.

I want to thank all PhD students and interns at our department for the breaks, the discussions we had including the last gossips from the Observant, and I would like to highlight our cozy PhD Christmas dinner (thanks interns for occupying your office). I would also like to thank my roommate *Maya Schulpen* for having a nice and calm companion during my PhD who was always prepared to listen to me when I needed it. I would also specifically like to thank *Ivette, Denise, Noreen, Louise, Linda, Sophie* and *Jacqueline* for their support during my PhD. *Catherine Mbakwa* for convincing and empowering me to sing Christmas Medleys together with her in front of the entire department of Epidemiology at the annual Christmas dinner. I am fascinated by your strong-mindedness and persuasive skills to pull me behind my computer screen, while I was hiding there.

I would like to thank my paranymphs. I am grateful for *Annaleen Koole* for supporting me in all my endeavors and listening to pretty much all of the ideas and possible outcomes that precede these endeavors. Even when you doubt about my sanity for some of those thoughts, you would always keep on listening and give me time to express these thought out loud, which then allows me to see that most of the ideas will remain just thoughts and need no further action. Thank you for supporting me throughout my PhD and for sharing so many nice memories together. I am also grateful for *Pierre Thomas*, "too French to be polite". I love how our conversations help me to lighten up through difficult times, either through your unpolished wit and quirkiness but also through more serious conversations. You know exactly how to levitate between these states of comfort and if none of these worked out you would offer me to have food at your place as a last resort, you must have been pleasantly surprised that this rather primitive method actually turned out well every single time!

I would also like to thank *Carolin Sehlbach* for being a support throughout my PhD, talking with you is always very insightful for me because of your sometimes different view on things that widen my horizon. I would like to thank *Elina Miteniece* for keeping me in check with myself when I'm doing too many things, probably because you understand yourself what it is like to having the urge to put 48 hours in one day.

I am grateful *to Johan Lugtenburg*, who is now emeritus Professor at Leiden University and still going strong for conducting research. I want to thank you for mentoring me while I was still doing my masters, broadening my knowledge during one-on-one discussions and suggesting me books and news articles. You really shaped my future with transcending your wisdom to me. *Jan & José de Pauw*, for believing in me when I did not believe in myself and shaping my professional and personal outlook. I want to *thank Ron de Kloet, Erno Vreugdenhil and Carla Verissimo* for supporting my exchange abroad during my studies in Leiden, owing to the mobility, I saw great things, learned and experienced new cultures, and met people (a lot of epidemiologists in Sweden) who changed my life for the better. I would like to thank *Clio Stronk* and *Susanne Bartels* for their unconditional support and motivational speeches during my time in Leiden but also throughout my PhD in Maastricht.

*Oma Annie Stroink*, ik heb met jou en *oma Annie* en *opa Jan Elands*, altijd een erg speciale band gehad, ik kon altijd bij jullie terecht en jullie zijn een vitale steun geweest voor de ontwikkeling in de persoon die ik nu ben. Opa Elands jouw geschiedenis is ook een aanleiding geweest voor mijn interesse in dit PhD project, omdat jij en je zus tijdens de Tweede Wereldoorlog wegens ondervoeding ondergebracht zijn geweest in een adoptiegezin in Engeland, leeft deze geschiedenis voor mij. Tot op heden zijn we nog steeds in contact met deze familie uit Engeland.

*Vader* en *moeder* bedankt voor jullie onvoorwaardelijke steun, liefde, en alle adviezen van jullie over allerhande zaken waarin ik in het leven tegenaan ben gelopen of wanneer ik weer eens een opbeurend gesprek nodig had. *Shirley*, mijn allerbeste vriendin en zusje. Ik kan altijd bij jouw terecht voor onbevooroordeeld advies, waarin ik jou vaak vraag: "Wat denk je dat ik het beste kan doen in deze situatie?", hopende op een strategisch en analytisch doordacht advies, iets wat jou op het lijf geschreven is. Maar dan laat je je van een andere kant zien en drukt me op het hart om ook naar mijn gevoel te luisteren. Jij hebt een verfijnde neus

voor mijn ontwikkelpunten en voelt feilloos aan waar ik mee worstel, zelfs voordat ik mij er zelf van bewust ben. Jij hebt me geleerd mijn eigen keuzes te maken mede op basis van mijn gevoel en de verantwoordelijkheid hiervoor te nemen, ook al zijn de consequenties soms onprettig. Ik begrijp goed waarom jij dingen gedaan krijgt waar je enorm veel energie uit put, omdat je naar eigen zegge eigen-wijs bent en je eigen pad volgt, daar is moed voor nodig! Je fascineert en inspireert mij!

Trevi bedankt dat je altijd voor mij klaar staat en bereid bent om mij met je ondernemersgeest en intellect op allerhande zaken te adviseren."

*Lukas*, jij bent mijn onvoorwaardelijke steun en toeverlaat geweest tijdens mijn PhD. Vooral tijdens de laatste periode waarin ik heel erg op mijzelf gericht was en ik niet altijd de meest sociale persoon was. Je bent ontzettend loyaal, betrouwbaar, zorgzaam en staat altijd voor mij klaar. Hoewel je vaak een duidelijke weloverwogen mening hebt, moedig je me tegelijkertijd aan het doel te volgen dat ik voor ogen heb, ook als je dit doel niet begrijpt. Ik waardeer deze eigenschap ontzettend, hier heb ik veel kracht uit kunnen putten!

Volgens ons niemand te vergeten.

Taller people have a greater chance of developing different types of cancer than shorter people. Although the reason for this is still unclear, a common mechanism might be at play. This thesis is aimed at obtaining a deeper understanding of how adult-attained height is associated with cancer risk later in life. As the variation in adult height is determined by genetic and environmental factors during the first 20 years of life, genetic and early life environmental factors influencing growth might also be related to cancer risk in later life. This thesis describes how adult-attained height and early life energy restriction during critical time windows of growth are associated with postmenopausal breast cancer in the Netherlands, currently inhabiting the tallest population in the world, by making use of the Netherlands Cohort Study. In addition, a systematic literature review and meta-regression analysis on early-life energy restriction and cancer risk in humans is described. Finally, the thesis contains a systematic selection approach for genetic variants to identify mechanisms underlying the link between height and post-menopausal breast and colorectal cancer.